MECHANISTIC STUDIES OF CLASS II LANTHIPEPTIDE SYNTHETASES AND YEAST SURFACE DISPLAY OF LANTHIPEPTIDE LEADER PEPTIDES

BY

YI YU

DISSERTATION
Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Biochemistry
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2015

Urbana, Illinois

Doctoral Committee:

Professor Wilfred A. van der Donk, Chair
Professor David M. Kranz
Professor John A. Gerlt
Professor Paul J. Hergenrother

ABSTRACT

Lanthipeptides are natural products that belong to the family of ribosomally synthesized and posttranslationally modified peptides (RiPPs). They contain the characteristic lanthionine (Lan) or methyllanthionine (MeLan) structures that contribute to their diverse biological activities. ProcM is a promiscuous bifunctional synthetase that catalyzes both dehydration and cyclization of the lanthipeptides. Its 30 substrate peptides (ProcAs) have a high level of conservation in the N-terminal leader region and hypervariability in the C-terminal core region. A chimeric peptide with a ProcA leader peptide and LctA (the precursor peptide for lacticin 481) core peptide was constructed and co-expressed with ProcM in *Escherichia coli*. ProcM dehydrated this chimeric peptide up to 5-fold (WT lacticin 481 is dehydrated 4-fold), and also catalyzed the formation of 2~3 rings (WT lacticin 481 has 3 rings). The modified peptide was digested with protease to remove the leader peptide, and the resulting product was demonstrated to have antimicrobial activity against *Lactococcus lactis HP*. These results showed the promiscuity of ProcM and suggest it may be used to generate cyclic peptide libraries from substrates with a ProcA leader peptide.

The product specificity and the regulation of product formation of ProcM were also studied. Despite its structurally very diverse substrate set, high selectivity of product formation from each substrate was observed. The effects of mutation of the conserved residues in the active site, including three zinc ligands (Cys924, Cys970, and Cys971), the proposed active site acid/base His859, and Asp804 that is hydrogen-bonded to His859 were investigated. Mutation of the zinc ligands to alanines or the unique zinc ligand Cys971 to histidine resulted in a decrease of the cyclization rate, especially for the second cyclization of ProcA1.1, ProcA2.8 and ProcA3.3. In the case of ProcA3.3, these mutations also altered the regioselectivity of ProcM and generated

a new major product. The H859A mutation completely deactivated the cyclization activity of the enzyme, while the D804N mutation also decreased the cyclization rate but maintained the initial regioselectivity for ProcA3.3. ProcM was not able to correct the ring topology of incorrectly cyclized intermediates and products, suggesting that thermodynamic control is not operational. Instead, the high regioselectivity of product formation appears to be governed by the selectivity of the initially formed ring.

By studying truncated enzymes of ProcM, LctM, and CylM, I showed that the N-terminus of these LanMs catalyzed the dehydration and the C-terminal domain catalyzed the cyclization. These two domains were active on their own, indicating that they are relatively independent. The N-terminus of LctM completely dehydrated its substrate LctA with efficiency similar to WT LctM. The N-terminus of CylM also dehydrated CylLs completely. The N-terminus of ProcM catalyzed one dehydration efficiently in ProcAs, but in the presence of the C-terminus, the dehydration reactions were complete. These results may suggest a different mode of catalysis for these LanMs.

The precursor peptides of lanthipeptides are ribosomally translated and are composed of an N-terminal leader peptide and a C-terminal core peptide. Sequence alignments of class II lanthipeptide leader peptides showed that they are typically rich in aspartates and glutamates. The leader peptide is believed to be recognized by the biosynthetic enzymes and guide the process of the post-translational modifications. Moreover, it is proposed that the binding of leader peptides to the enzymes shifts the equilibrium of the inactive and active enzymes to the latter form. Efforts were made to obtain an LctA leader peptide with better binding affinity to LctM for the following reasons: 1. Increased binding affinity of LctA to LctM might result in improved catalytic efficiency, and 2. Increased binding affinity might facilitate finally obtaining

a crystal structure of a LanM with the substrate bound to it. Yeast surface display was selected to engineer the LctA leader for better binding affinity to the LctM enzyme. Two rounds of sorting resulted in several mutations in the leader region, including three consistent aspartate mutations. Using fluorescence polarization assay, it was confirmed that the binding affinity of the mutant leader peptide was about 5-fold stronger than the WT leader peptide. Furthermore, LctA containing the engineered leader peptide was shown to be a more efficient substrate for LctM.

To my beloved husband,
my parents,
and my unborn daughter.

ACKNOWLEDGEMENT

I would like to gratefully and sincerely thank my advisor Dr. Wilfred van der Donk for his guidance, understanding, patience, and support throughout the six years of my graduate studies. His impeccably rigorous scientific thinking, highly responsible work attitude, well-balanced attention to his family as well as his sincere love and patience for his students have set the best example for me both in work and in life. He provided me with support when I took risks in the research projects, encouragement when I went thought difficulties in both work and life, patience when I made different kinds of mistakes, and guidance when I was in great need of them. For everything you've done for me, Wilfred, I couldn't thank you enough.

I would like to thank several important people in my life. My parents always provide me with trust and support for almost everything in my life. Their unconditional love have encouraged me to pursue my dreams even when the road to my goal is not always smooth, and their understanding keep me motivated during the long journey of graduate school.

I would not forget the tough decision my husband Yang made when his whole group moved to California. Meeting him was the luckiest thing that has ever happened to me. With his companion, I enjoyed my graduate school so much more, and I am looking forward to everything that is waiting for us in the future. A special thank is given to our unborn daughter. She went through the last part of graduate school with me, and made this thesis writing so much fun.

I would like to thank all my friends in this small quiet town. We shared joy and sadness, dreams about future, immature opinions, as well as the precious memories. As they say, you will miss the worst days you have with your young friends even in the best days you have on your own.

I would also like to thank the rest of my committee members, Professor David Kranz, Professor John Gerlt, and Professor Paul Hergenrother, for their continued help and support. I

would like to thank all of the members of the van der Donk research group, especially Dr. Yanxiang Shi and Dr. Trent Oman for training me when I joined the lab, Dr. Weixin Tang and Dr. Xiao Yang for growing together with me, Lindsay Johnstone, Dr. Ran Zhang, Xiling Zhao for creating a great environment to work in. I would also like to thank Dr. Shihui Dong and Jonathan Chekan from Professor Satish Nair' s lab, Dr. Adam Kelvin from Professor David Kranz lab, Dr. Babara Pilas from the Flow Cytometry Facility and Dr. Zhong Li from the Metabolomics Center for the help with experiments.

TABLE OF CONTENTS

CHAPTER I: INTRODUCTION-THE BIOSYNTHETIC PATHWAY AND MODE OF ACTION OF LANTHIPEPTIDES*

1.1 OVERVIEW

Natural products have been a proven source for drug and drug leads discovery, especially in the area of antimicrobial and anticancer drugs. In the 20th century, many classes of natural products have been identified including four particularly prevalent groups: terpenoids, alkaloids, polyketides, and non-ribosomal peptides.[1] The genome sequencing efforts of the past decade have illustrated that ribosomally synthesized and posttranslationally modified peptides (RiPPs) are much more prevalent than previously anticipated.[1] RiPPs are encoded in the genomes of organisms in all domains of life, their structural diversity is vast, and their biological activities are equally diverse. Most RiPPs are biosynthesized from a larger precursor peptide that consists of a core peptide that is converted to the final product and an N- or C-terminal extension called the leader or follower peptide, respectively. These extensions guide many, but not always all, of the posttranslational modifications that take place in the core peptide.[2]

---

*Reproduced in part with permission from Yu, Y., Zhang, Q., and van der Donk, W. A. (2013) Insights into the evolution of lanthipeptide biosynthesis, *Protein Sci. 22*, 1478-1489.

RiPPs are typically macrocyclic, similar to the high incidence of cyclization in non-ribosomal peptides. Cyclization can offer several advantages such as increased metabolic stability, improved cellular uptake, and preorganization to recognize cellular targets. Lanthionine-containing peptides, or lanthipeptides, are the largest group of RiPPs based on the frequency of their biosynthetic gene clusters in the currently available genomes. Over 90 members of the lanthipeptide family have been identified to date, most with distinct ring topologies.[3] This thesis describes studies designed to gain insight into the lanthipeptide biosynthesis process and lanthipeptide engineering.

1.2 THE STRUCTURES OF LANTHIPEPTIDES

As mentioned above, lanthipeptides are named after the characteristic lanthionine structures. The lanthionine structure, abbreviated as Lan, consists of two alanine residues that are linked through a thioether that connects their b-carbons (Figure 1.1). Many lanthipeptides also contain methyllanthionines, abbreviated as MeLan, which carry an additional methyl group on one of the β-carbons (Figure 1.1). Biosynthetically, lanthionines and methyllanthionines originate from Ser and Thr residues that are first dehydrated to generate dehydroalanine (Dha) and dehydrobutyrine (Dhb) residues, respectively (Figure 1.1).[3] Subsequently, the thiol of a Cys is added across the carbon–carbon double bond of these dehydroamino acids in a Michael-type addition to produce the Lan and MeLan structures, respectively (Figure 1.1). Since typical lanthipeptide substrate peptides have multiple Ser, Thr, and Cys residues, the final products are polycyclic and display remarkable diversity in structure as illustrated by a select set of lanthipeptides depicted in Figure 1.2.

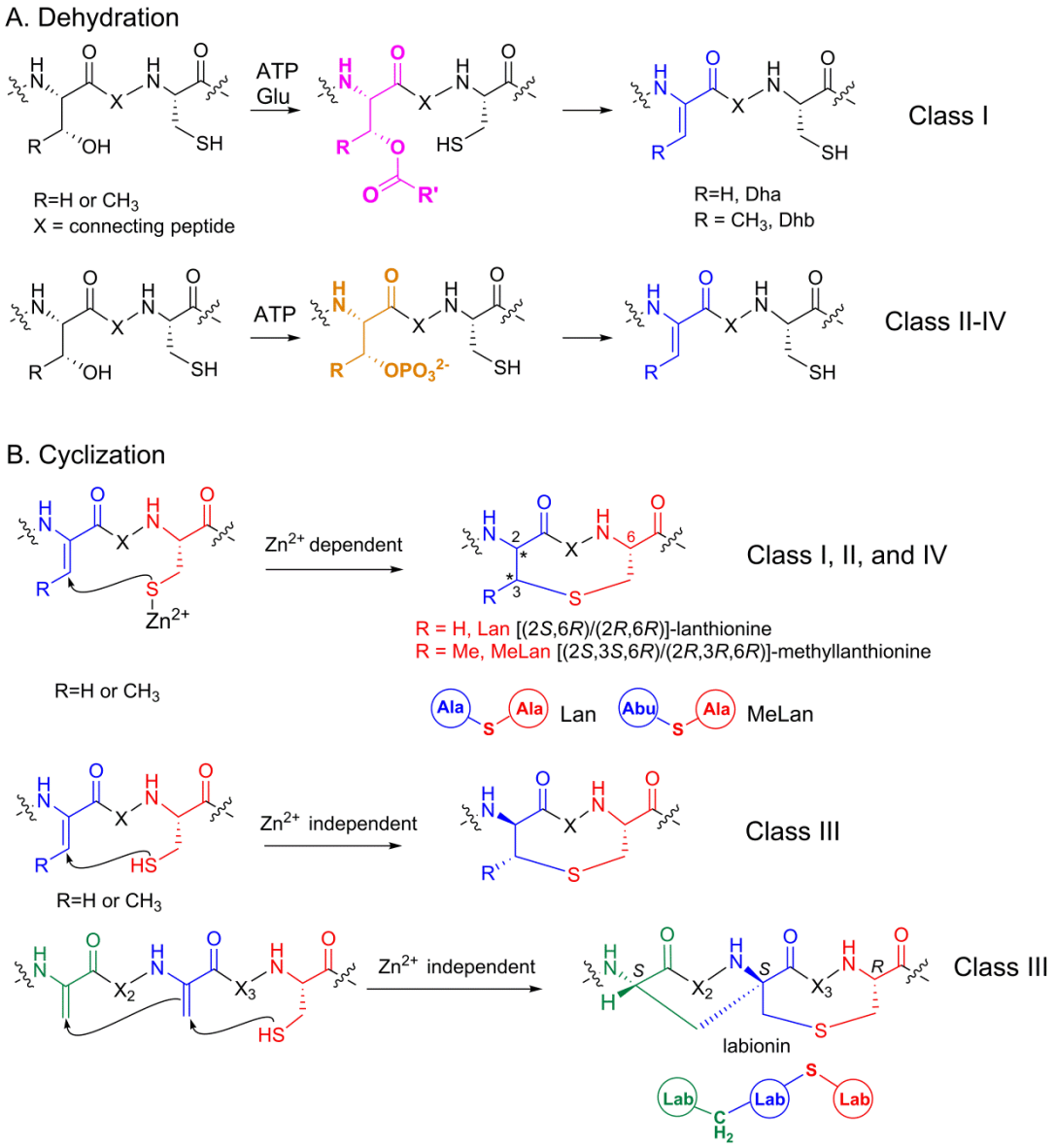

Figure 1.1. A. The dehydration step in lanthionine formation involves activation of the side chain hydroxyl group of Ser and Thr to generate dehydroamino acids (blue). For the class I enzyme NisB involved in nisin biosynthesis, activation involves glutamylation (magenta). The ester bond is likey made with the α carboxylate of glutamate. For class II-IV lantibiotics, activation of the hydroxyl group involves phosphorylation (orange) and elimination. **B.** The cyclization step in lanthipeptide biosynthesis also comes in different flavors. For class I, II, and IV lanthipeptides it is believed that cyclization involves activation of the Cys nucleophile (red) by an active site Zn^{2+}. For class III lanthipeptides, the Zn^{2+} is not present in the cyclase and

thiol activation is achieved by an alternative, currently unknown mechanism. Class III cyclization is also unusual in that it results either in the formation of (methyl)lanthionines or the formation of labionins, which contain a carbon-carbon crosslink formed from one Cys (red) and two dehydroalanines (blue and green).[4] An example is labyrinthopeptin shown in Figure 1.2.2. Abu, 2-aminobutyric acid; Lab, labionin.

Currently known lanthipeptides have a range of different activities including antimicrobial (called lantibiotics),[5] morphogenetic,[6] antiviral,[7] and antiallodynic.[4] The longest known family member, nisin (Figure 1.2), has been used for more than 50 years as a food preservative to combat food-borne pathogens.[8] In addition, several other lanthipeptides are currently under development for various therapeutic applications including a semisynthetic analog of actagardine,[9] the chlorinated lantibiotic microbisporicin,[10, 11] duramycin,[12] and labyrinthopeptin A2 (Figure 1.2).[4]

Diverse posttranslational modifications besides Lan/MeLan are also present in lanthipeptide family (Figure 1.2), including β-hydroxylated aspartate and lysinoalanine (in cinnamycin), aminovinylcysteine (in mersacidin and others), D-alanine (in lacticin 3147 and lactocin S), and several α-ketoamide and α-hydroxyamide N-terminal caps (in Pep5, lactocin S, and others). Many of these modifications are considered to improve peptide stability through protection from proteolysis. In the case of cinnamycin, the β-hydroxylated-Asp15 residue is believed to contribute to the interaction with the glycerophosphoethanolamine head group of its target phosphatidylethanoamine (PE), which is a component of the bacterial membrane.[13]

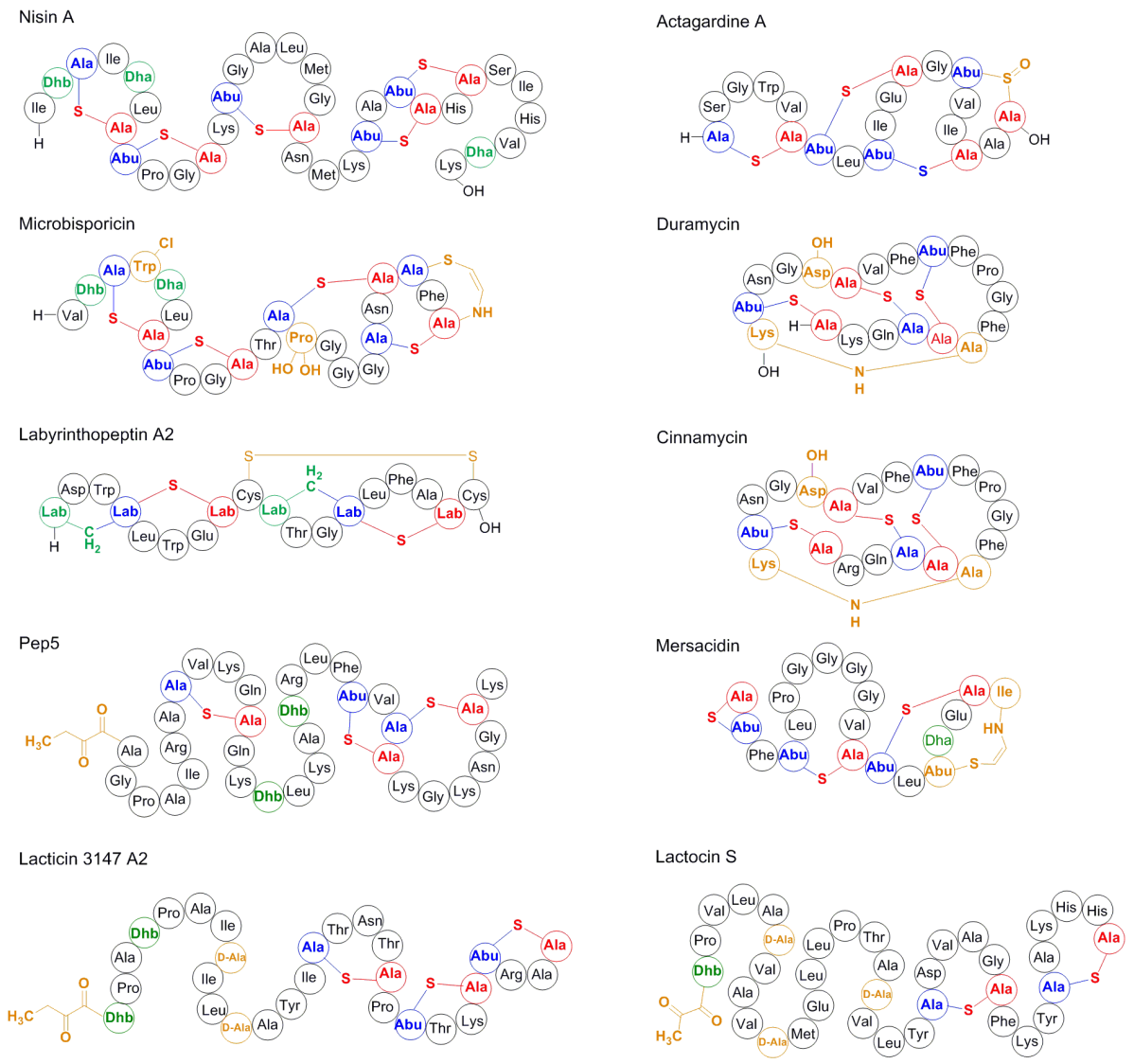

Figure 1.2. Structures of a representative set of lanthipeptides, demonstrating the vast diversity of ring topologies. The same short hand notation is used that is shown in Figure 1.1. (Me)Lan segments derived from Cys are shown in red and originating from Ser/Thr are shown in blue. Additional posttranslational modifications are shown in orange. Dehydroamino acids are shown in green.

1.3 BIOSYNTHESIS OF LANTHIPEPTIDES

The lanthipeptide precursor peptides (generically termed LanAs) contain an N-terminal leader region in addition to a C-terminal core peptide that is processed to the mature compound. The general route for Lan/MeLan synthesis involves the dehydration of select serine and threonine residues to dehydroalanine (Dha) and dehydrobutyrine (Dhb) residues, respectively, followed by intramolecular conjugate additions of cysteines to the dehydro amino acids, resulting in extensively cross-linked polycyclic structures. After complete modification of the precursor peptide, the N-terminal leader sequence is removed by a protease to release the mature peptide (Figure 1.3).

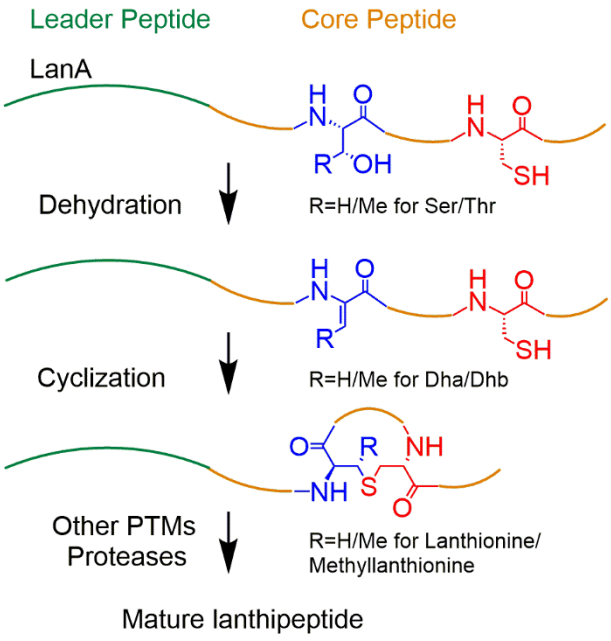

Figure 1.3. Overview of class II lanthipeptide biosynthesis. The canonical post-translational modifications are shown.

Although all lanthipeptides are made by dehydration of Ser and Thr residues followed by the addition of Cys residues to the resulting dehydroamino acids, the catalytic machinery that carries out these reactions is remarkably diverse. At present, four different enzymatic processes to lanthionines have been identified, some of which are evolutionarily related (Figure 1.4)[3]. Class I lanthipeptides are synthesized by two separate enzymes. The dehydration reactions are catalyzed by a dehydratase LanB, and the cyclization reactions are catalyzed by a cyclase LanC. For class II lanthipeptides, both the dehydration and cyclization reactions are carried out by a bifunctional enzyme LanM, which contains an N-terminal dehydratase domain that has no homology with LanB and a C-terminal LanC-like cyclase domain. Chapter II and III describe studies designed to understand one particular class II synthetase called ProcM, and chapter IV describe studies on the individual domains of several class II synthetases. The synthetases for class III lanthipeptides (LanKC) are trifunctional enzymes with an N-terminal lyase domain, a central kinase domain and a C-terminal putative cyclase, which does not contain the conserved zinc ligand residues. The synthetases for class IV lanthipeptides (LanL) share a similar scaffold with LanKC, except that the C-terminal cyclase domain is homologous to LanC.

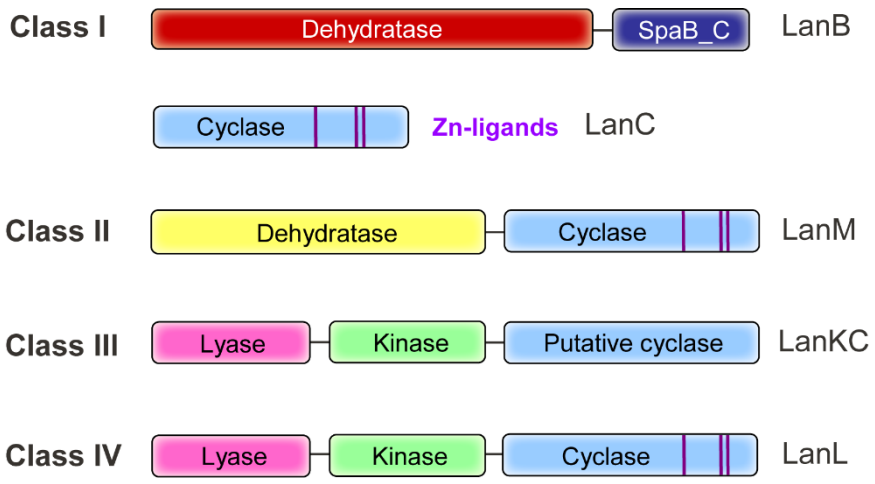

Figure 1.4. Schematic representation of the four different classes of lanthionine synthetases. Class I systems consist of dedicated dehydratases (generically called LanB) and cyclases (generically called LanC). Class II enzymes are bifunctional, containing an N-terminal dehydratase domain and a C-terminal cyclase domain that has homology with the LanC proteins including the Zn-ligands (purple lines). Class III and IV enzymes contain a central kinase domain and an N-terminal lyase domain. The C-termini of these synthetases differ, with class III enzymes containing a cyclization domain that has homology with LanC proteins but that lacks the Zn-ligands, whereas class IV enzymes have the canonical LanC-like cyclization domain.

1.3.1 Class I lanthipeptide biosynthesis

For nisin and other class I lanthipeptides, the dehydration reaction is carried out by LanB dehydratases and the cyclization by LanC cyclases. The LanB enzymes are large, about 120 kDa, and do not have homology with any other characterized proteins in the databases. NisB,

the LanB dehydratase involved in nisin biosynthesis, requires glutamyl-$tRNA^{Glu}$ to carry out the dehydration of eight Ser/Thr residues in the precursor peptide.[14, 15] In this process, the hydroxyl groups of Ser and Thr are first converted to glutamyl esters, which are then eliminated to form the carbon–carbon double bonds (Figure 1.1). The transient glutamylation constitutes a novel posttranslational modification of a peptide or protein; the only other known examples of glutamylation of Thr occur not on a peptide or protein substrate but rather on the enzymes glutaminase-asparaginase and γ-glutamyltranspeptidase as a catalytic acyl enzyme intermediate during the hydrolysis of glutamine and glutathione, respectively.[16, 17]

The addition of the thiol of Cys to Dha and Dhb catalyzed by NisC, the cyclase involved in nisin biosynthesis, is believed to involve an active site Zn^{2+} ion that was identified by both spectroscopic and crystallographic studies.[18, 19] Analogous to other enzymes that activate thiol nucleophiles such as farnesyl transferase,[20] the Zn^{2+} ion is believed to activate the Cys thiols in the substrate peptide. Not only is the mechanism of catalysis similar, NisC also has structural homology with farnesyl transferase, sharing the same α,α-barrel toroidal fold, with a Zn^{2+} at the top of the barrel.[19] Interestingly, whereas LanB enzymes have no sequence homologs other than other putative dehydratases, LanC-like (LanCL) proteins of unknown function are also found in a wide variety of higher organisms including plants,[21] insects, and mammals.[22, 23]

1.3.2 Class II lanthipeptide biosynthesis

The class II enzymes, generically called LanM, first phosphorylate Ser/Thr residues in the substrate peptides, followed by elimination of the phosphate to generate the dehydroamino acids (Figure 1.1).[24] This net dehydration activity is located in the N-terminal domain of LanM proteins, which are about 110–120 kDa in size. Like the LanB proteins, the dehydratase

domains of LanM proteins have no clear sequence homologs in the protein databases other than other class II lanthionine synthetases, but the C-terminal cyclase domains of LanM proteins have clear sequence homology with the class I LanC cyclase enzymes, including the conserved Zn^{2+} ligands (Figure 1.4). Chapter III describes studies focused on the cyclization process of the class II synthetase ProcM.

1.3.3 Class III and class IV lanthipeptide biosynthesis

Class III and IV lanthionine synthetases, generically called LanKC and LanL, respectively, display an interesting domain architecture. These enzymes contain a central domain with clear sequence homology with Ser/Thr protein kinases (Figure 1.4).[25] Indeed both classes of enzymes have been shown to phosphorylate the Ser and Thr residues in the substrate peptide that are destined to be dehydrated,[26, 27] similarly to the catalytic strategy employed by the class II LanM enzymes. However, the elimination of the phosphate group to form the alkenes appears to have evolved differently because class III and IV enzymes contain a phosphothreonine (pThr) lyase domain at their N-termini that is not found in LanM enzymes,[26, 28] at least not at the sequence level; the possibility of structural homology cannot be ruled out in the absence of crystal structures. Phosphothreonine lyases are employed as effector proteins in various pathogens like *Shigella* to attenuate part of the immune response of the host by eliminating a phosphate group from a phosphorylated Thr residue in MAP kinases.[29] Whereas the N-terminal lyase and central kinase domains are very similar in class III and class IV enzymes, the C-termini of these enzymes differ. Class IV proteins again have a canonical LanC cyclase domain including the conserved Zn^{2+} ligands,[26] but class III enzymes contain a C-terminal domain with sequence homology to LanC but lacking the Zn^{2+} ligands.[30] Indeed,

analysis of one class III enzyme, AciKC, showed that the C-terminal domain is essential for cyclization and that the enzyme did not contain Zn nor any other metals. The cyclization domains of class III and IV enzymes also differ in the reactions they catalyze. Whereas the only class IV enzyme that has been analyzed thus far generates (methyl)-lanthionines,[26] the class III enzymes form either (methyl)lanthionines[25, 31] or labionins.[4, 30] The latter structures are believed to be formed by initial attack of a Cys residue onto a dehydroamino acid, resulting in an enolate. Instead of protonation of the enolate, which would give a (Me)Lan structure, it presumably attacks another dehydroamino acid, resulting in a carbocyclic structure (Figure 1.1).[22]

1.4 THE EVOLUTION OF LANTHIPEPTIDE SYNTHETASES

The high frequency of their occurrence in genomes as well as the emergence of four different routes to lanthipeptides likely reflects the ease by which complex polycyclic compounds can be generated by two relatively easy chemical reactions, water elimination from Ser/Thr and Michael type addition of Cys to dehydroamino acids. Enzymes involved in secondary metabolism often have descended from proteins involved in primary metabolism.[32] The current knowledge of the mechanisms and structures of lanthipeptide synthetases, and their sequence homology with known proteins, suggests that the modern lanthipeptide synthetases may have evolved from stand-alone posttranslational modification proteins.

For class III and IV this is most obvious, as fusion of protein kinase and pThr lyase domains created a dehydratase. Whereas most protein kinases and pThr lyases have evolved to be highly specific with respect to their substrates, the respective domains of class III and IV

lanthipeptide synthetases appear to have maintained the low substrate specificity of primitive progenitors, but they have become dependent on the leader peptides for activation. For class II LanM enzymes, it remains to be established whether phosphorylation and elimination occur in the same or different active sites. The origin of the class I LanB dehydratases is less clear. Presumably, their ancestor was involved in utilization of glutamyl-$tRNA^{Glu}$, but for what purpose is currently not clear.

It is interesting that whereas the dehydration enzymes apparently have independently evolved at least three times, the cyclization domains all have clear sequence homology. The closer relation of the four cyclization domains/enzymes is somewhat surprising since Michael-type addition of Cys residues to dehydroamino acids is a relatively facile process that even takes place readily in the absence of any enzyme.[33-37] Hence, a priori one might have expected greater divergence in the enzymes that catalyze the cyclization reaction. For the Zn-dependent enzymes, presumably their ancestor had a different activity involving a Zn^{2+} and they were recruited because they accelerated the Michael addition process and/or they fortuitously favored the formation of products with a ring topology that provided an evolutionary beneficial activity. Interestingly, the phylogenetic analysis shows that the eukaryotic LanCL proteins are closer related to the cyclization domain of class II LanM proteins than the standalone LanC enzymes (Figure 1.5).

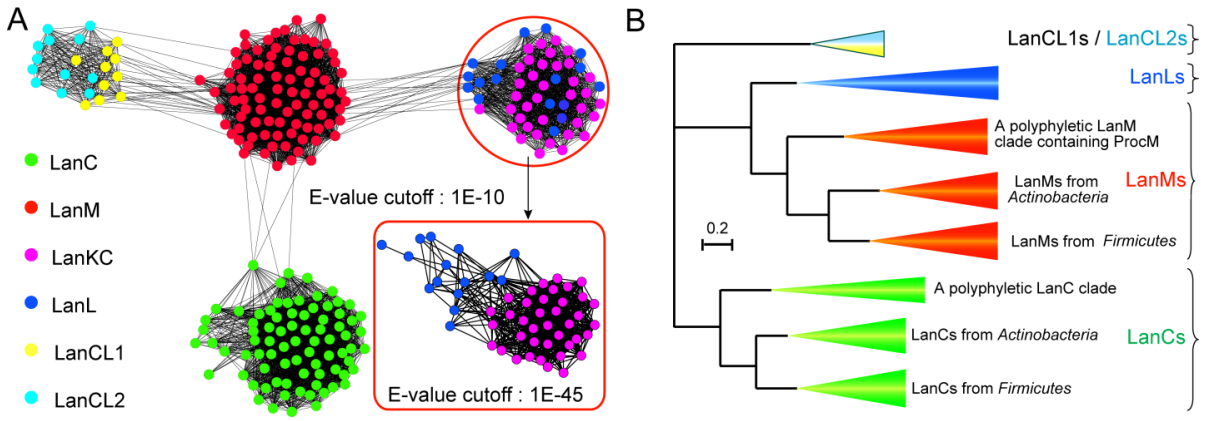

Figure 1.5. A. Cytoscape diagram showing the various cyclase-containing enzymes. In green are the class I LanC clade, in red the class II LanM clade, in magenta the class III LanKC enzymes, in blue the class IV LanL proteins, and in yellow and turquoise, the mammalian LanCL1 and LanCL2 proteins, respectively. **B.** Bayesian MCMC phylogeny of the class I, II, and IV cyclization enzymes as well as the LanCL proteins found in *Eukarya*. The ProcM-like enzymes are found in a separate subclade in the LanM clade.[38] Color scheme is the same as in panel A. Figure generated by Dr. Qi Zhang.

1.5 LANTHIPEPTIDE PRECURSOR PEPTIDES AND THEIR POSSIBLE EVOLUTIONARY HISTORY

The most challenging task for the lanthipeptide biosynthetic enzymes is control over the ring topology of their products. As can be seen from the structures of a small subset of currently known lanthipeptides in Figure 1.2, each cyclase enzyme generates a series of rings that all have different sequences and sizes. To provide an idea of this challenge, for nisin the dehydrated substrate peptide contains five Cys and eight dehydroamino acids (Figure 1.6). Thus, the number of isomers differing in ring topology that can be generated by a cyclization process lacking any selectivity would be 6720.[39] When also considering all of the possible

stereoisomers that can be formed, this number would be at least 8.6×10^5. However, NisC generates a single product out of all these potential structures. How the cyclization enzymes achieve these feats, which would be very challenging for a chemist, is still an open question, but some insights have recently started to emerge. One possibility is that the fully dehydrated intermediate drawn in Figure 1.6 actually is never formed. For instance, it is possible that the dehydratase and cyclase may be alternating in modifying the substrate peptide such that cyclization commences well before all dehydrations have taken place.[40-42] This scenario would decrease the number of electrophiles that each Cys has to distinguish. Indeed, directionality has been observed for several lanthipeptide synthetases.[30, 40, 41, 43]

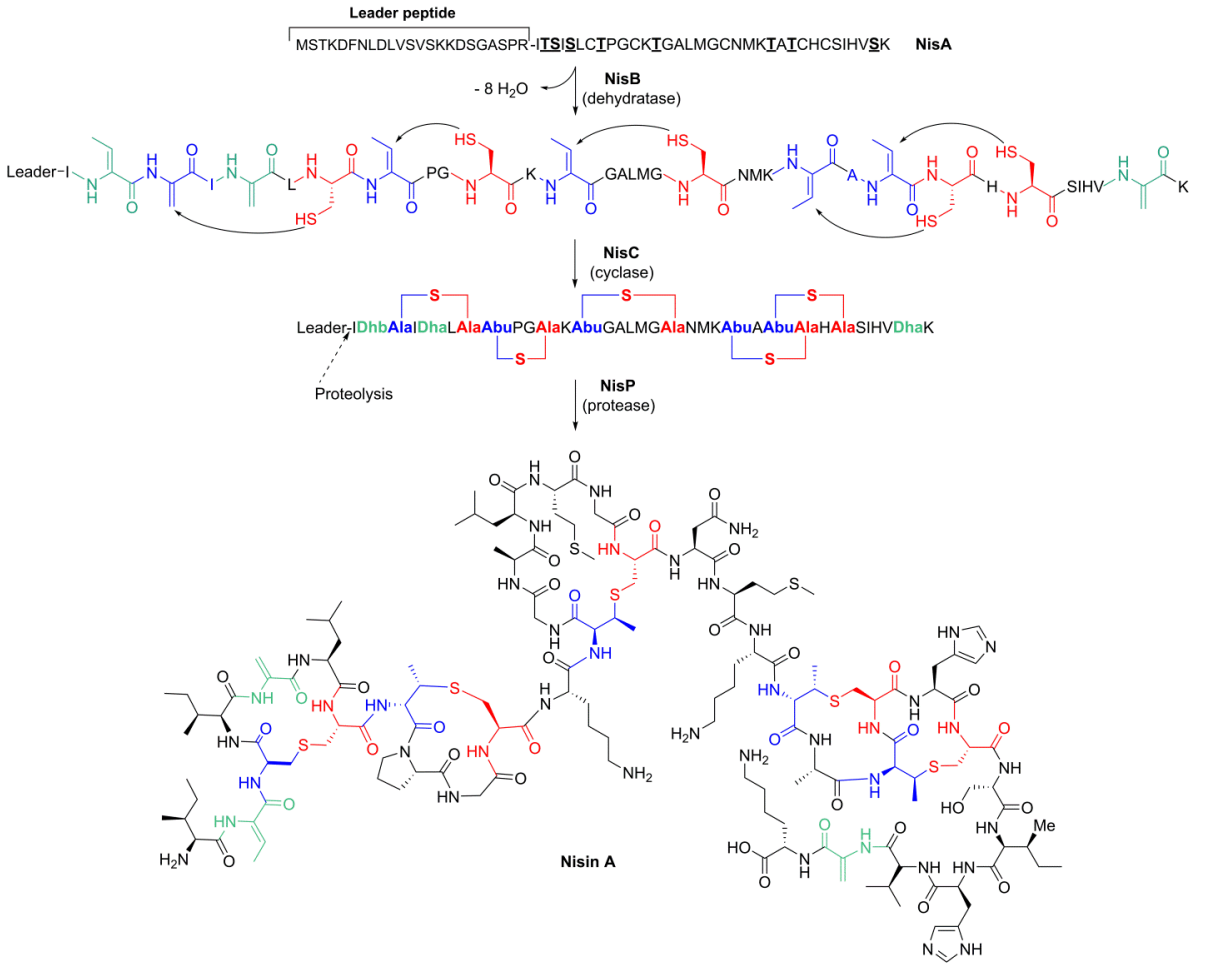

Figure 1.6. Biosynthesis of nisin A.

1.6 THE ROLE OF LANTHIPEPTIDE LEADER PEPTIDES

For almost all RiPPs, the precursor peptides contain an N-terminal leader peptide in addition to the C-terminal core peptide that is processed to the mature compound. There are several proposed roles for the leader peptides including guiding the biosynthesis process and providing self-immunity of the producing strain. The role most commonly proposed for the leader peptides is that of a secretion signal. However, the leader peptides of most RiPPs have no homology with the peptides of the typical secretory (Sec) pathways in archaea, the twin-arginine translocation pathways that are used in bacteria, and protein translocation pathways in plants.[2] Another frequently postulated role for the leader peptide is that of a recognition motif for the post-translational modification (PTM) enzymes in the biosynthesis pathways. From a natural product engineering perspective, this function is of great interest as it may allow generation of analogs of RiPPs simply by attaching core peptide variants or even very different peptides to the leader peptides. A related proposed function for the leader peptide is that of a *cis*-acting chaperone to actively assist the PTM process. Other possible functions for the leader peptides include providing assistance in the proper folding of the precursor peptide, stabilizing the precursor against degradation, and keeping the peptide inactive during biosynthesis inside the host cell. Support for almost all of these roles has been reported, but the function of the leader peptide differs for different classes of RiPPs and sometimes even within a certain natural product group.

For class I lanthipeptides, the leader peptides are typically 25 amino acids in length and are rich in aspartate residues.[2] NMR studies on the fully modified nisin with the leader peptide attached showed no interactions between the leader peptide and the nisin molecule in aqueous

solution and in micelles.[44] Moreover, none of the investigated lanthipeptide leader peptide showed any secondary structures in aqueous solution, although they displayed α-helical conformation in trifluoroethanol,[45] and computational analysis based on the amino acid sequence predicted helical structures. Together, these observations argue against the hypothesis that the leader peptides act as *cis*-acting chaperone for folding. Leader peptides attached to the modified core peptides reduce or abolish the antimicrobial activities, suggesting a role of leader peptides in self-immunity of the producing strain. Moreover, the leader peptide of nisin was fused to nonlanthipeptide cargo peptides and was successfully processed by the ATP-binding cassette (ABC) transporter NisT,[46] the dehydratase NisB,[47, 48], and the cyclase NisC.[19] Crystallographic studies showed that NisC is made up of an α,α-barrel that contains the active site and a separate domain proposed to contain the leader peptide binding site. More directly, a co-crystal structure of NisB and NisA provided the first structural insights into the leader peptide binding of lanthipeptide synthetases.[15]

For class II lanthipeptides, the leader peptides are also typically rich in negatively charged aspartate and glutamate residues and usually end in a double glycine (Gly-Gly or Gly-Ala/Ser/Thr) motif, which is the protease cleavage site.[2] Like class I lanthipeptides, the leader peptides of class II lanthipeptides also participate in the self-immunity of the host cell by keeping the modified compound inactive inside the cell. Interestingly, *in vitro* studies on lacticin 481 have shown that the leader peptide is important for efficient dehydration and cyclization catalyzed by LctM, but is not essential.[49] Moreover, an LctM enzyme with the LctA leader peptide attached through a flexible linker was able to modify linear LctA core peptide without leader sequence.[50] These observations indicate that the binding of the leader peptide to

the enzyme might shift the equilibrium of inactive and active enzyme towards the latter form. However, for class II lanthipeptides synthetases, no crystal structures and co-crystal structures with the substrate bound have been reported so far. Further structural and mechanistic investigation of the functions of lanthipeptide leader peptides will provide deeper understanding of the biosynthesis of the lanthipeptides and may improve the opportunities to engineer lanthipeptides to tune their activities for specific purposes. Chapter V provides one potential strategy, the use of yeast surface display to select for tighter binding leader peptides.

1.7 POTENTIAL APPLICATIONS OF LANTHIPEPTIDE BIOENGINEERING

The increasing appearance of drug resistance of antibiotics makes the discovery of new antibiotics an urgent topic.[51] As mentioned above, the lantibiotic nisin has been widely used as a food preservative for half a century without causing significant drug resistance. Lantibiotics have been thought not to be prone to development of resistance because many of them target the pyrophosphate group of lipid II, which is not easy to be mutated and some of the lantibiotics have a dual mode of action to reduce the chance of generating resistance in host cell.[52] Thus, the lanthipeptides are great candidates for alternative drugs. Despite the promising antimicrobial activities displayed by many lanthipeptides, there are several factors preventing their wide therapeutic applications, including potency, stability, and solubility. In the past decades, an increased understanding of their biosynthesis pathway has facilitated the engineering efforts of lanthipeptides both *in vitro* and *in vivo*.[3] Besides the widely reported antimicrobial activities of lanthipeptides, the extensively cross linked structures makes them

one of the structurally most complex group in the RiPP family. The cyclic peptides are of great therapeutic interest due to their enhance stability and constrained structures. There are increasing studies on engineering cyclic peptide libraries to reach specific goals, including obtaining *de novo* activities.[53, 54] Lanthipeptides are potential candidates for the creation of cyclic peptide libraries for selection of improved properties as antibiotics as well as new activities such as disrupting protein-protein interactions of pharmaceutical interests.

The early manipulations of lanthipeptides were achieved using the producing strain or closely related heterologous hosts. In 2008, a complementation approach was reported combining random and site-saturation mutagenesis of *nisA* generated nisin mutants with a variable hinge region between the C and D rings, and several of these mutants demonstrated improved activities against pathogenic strains that were tested.[55] In 2011, our lab reported the production of diverse lanthipeptides including prochlorosins, haloduracin and nisin in *Escherichia coli* (*E. coli*), which greatly reduced the time and cost required for the production of lanthipeptides, significantly increased the yield, simplified the purification process, as well as allowed the incorporation of noncanonical amino acids.[56] However, there are no reports of the engineering of lanthipeptides for *de novo* functions, and the lanthipeptides family remains an under-explored resource for cyclic molecules of therapeutic interests.

1.8 REFERENCES

1. Arnison, P. G., Bibb, M. J., Bierbaum, G., Bowers, A. A., Bugni, T. S., Bulaj, G., Camarero, J. A., Campopiano, D. J., Challis, G. L., Clardy, J., Cotter, P. D., Craik, D. J., Dawson, M., Dittmann, E., Donadio, S., Dorrestein, P. C., Entian, K. D., Fischbach, M. A., Garavelli, J. S., Goransson, U., Gruber, C. W., Haft, D. H., Hemscheidt, T. K., Hertweck, C., Hill, C., Horswill, A. R., Jaspars, M., Kelly, W. L., Klinman, J. P., Kuipers, O. P., Link, A. J., Liu, W., Marahiel, M. A., Mitchell, D. A., Moll, G. N., Moore, B. S., Muller, R., Nair, S. K., Nes, I. F., Norris, G. E., Olivera, B. M., Onaka, H., Patchett, M. L., Piel, J., Reaney, M. J., Rebuffat, S., Ross, R. P., Sahl, H. G., Schmidt, E. W., Selsted, M. E., Severinov, K., Shen, B., Sivonen, K., Smith, L., Stein, T., Sussmuth, R. D., Tagg, J. R., Tang, G. L., Truman, A. W., Vederas, J. C., Walsh, C. T., Walton, J. D., Wenzel, S. C., Willey, J. M., and van der Donk, W. A. (2013) Ribosomally synthesized and post-translationally modified peptide natural products: overview and recommendations for a universal nomenclature, *Nat. Prod. Rep. 30*, 108-160.

2. Oman, T. J., and van der Donk, W. A. (2010) Follow the leader: the use of leader peptides to guide natural product biosynthesis, *Nat. Chem. Biol. 6*, 9-18.

3. Knerr, P. J., and van der Donk, W. A. (2012) Discovery, biosynthesis, and engineering of lantipeptides, *Annu. Rev. Biochem. 81*, 479-505.

4. Meindl, K., Schmiederer, T., Schneider, K., Reicke, A., Butz, D., Keller, S., Guhring, H., Vertesy, L., Wink, J., Hoffmann, H., Bronstrup, M., Sheldrick, G. M., and Sussmuth, R. D. (2010) Labyrinthopeptins: a new class of carbacyclic lantibiotics, *Angew. Chem., Int. Ed. Engl. 49*, 1151-1154.

5. Schnell, N., Entian, K.-D., Schneider, U., Götz, F., Zahner, H., Kellner, R., and Jung, G. (1988) Prepeptide sequence of epidermin, a ribosomally synthesized antibiotic with four sulphide-rings, *Nature 333*, 276-278.

6. Willey, J. M., and Gaskell, A. A. (2011) Morphogenetic signaling molecules of the streptomycetes, *Chem. Rev. 111*, 174-187.

7. Férir, G., Petrova, M. I., Andrei, G., Huskens, D., Hoorelbeke, B., Snoeck, R., Vanderleyden, J., Balzarini, J., Bartoschek, S., Brönstrup, M., Süssmuth, R. D., and Schols, D. (2013) The Lantibiotic Peptide Labyrinthopeptin A1 Demonstrates Broad Anti-HIV and Anti-HSV Activity with Potential for Microbicidal Applications, *PloS one 8*, e64010.

8. Lubelski, J., Rink, R., Khusainov, R., Moll, G. N., and Kuipers, O. P. (2008) Biosynthesis, immunity, regulation, mode of action and engineering of the model lantibiotic nisin, *Cell. Mol. Life Sci. 65*, 455-476.

9. Boakes, S., Ayala, T., Herman, M., Appleyard, A. N., Dawson, M. J., and Cortes, J.

(2012) Generation of an actagardine A variant library through saturation mutagenesis, *Appl. Microbiol. Biotechnol. 15*, 1509-1517.

10. Castiglione, F., Lazzarini, A., Carrano, L., Corti, E., Ciciliato, I., Gastaldo, L., Candiani, P., Losi, D., Marinelli, F., Selva, E., and Parenti, F. (2008) Determining the structure and mode of action of microbisporicin, a potent lantibiotic active against multiresistant pathogens, *Chem. Biol. 15*, 22-31.

11. Jabés, D., Brunati, C., Candiani, G., Riva, S., Romanó, G., and Donadio, S. (2011) Efficacy of the new lantibiotic NAI-107 in experimental infections induced by multidrug-resistant Gram-positive pathogens, *Antimicrob. Agents Chemother. 55*, 1671-1676.

12. Jones, A. M., and Helm, J. M. (2009) Emerging treatments in cystic fibrosis, *Drugs 69*, 1903-1910.

13. Wakamatsu, K., Choung, S. Y., Kobayashi, T., Inoue, K., Higashijima, T., and Miyazawa, T. (1990) Complex formation of peptide antibiotic Ro09-0198 with lysophosphatidylethanolamine: proton NMR analyses in dimethyl sulfoxide solution, *Biochemistry 29*, 113-118.

14. Garg, N., Salazar-Ocampo, L. M., and van der Donk, W. A. (2013) In vitro activity of the nisin dehydratase NisB, *Proc. Natl. Acad. Sci. U. S. A. 110*, 7258-7263.

15. Ortega, M. A., Hao, Y., Zhang, Q., Walker, M. C., van der Donk, W. A., and Nair, S. K. (2014) Structure and mechanism of the tRNA-dependent lantibiotic dehydratase NisB *Nature 517*, 509-512.

16. Ortlund, E., Lacount, M. W., Lewinski, K., and Lebioda, L. (2000) Reactions of Pseudomonas 7A glutaminase-asparaginase with diazo analogues of glutamine and asparagine result in unexpected covalent inhibitions and suggests an unusual catalytic triad Thr-Tyr-Glu, *Biochemistry 39*, 1199-1204.

17. Inoue, M., Hiratake, J., Suzuki, H., Kumagai, H., and Sakata, K. (2000) Identification of catalytic nucleophile of Escherichia coli gamma-glutamyltranspeptidase by gamma-monofluorophosphono derivative of glutamic acid: N-terminal thr-391 in small subunit is the nucleophile, *Biochemistry 39*, 7764-7771.

18. Okeley, N. M., Paul, M., Stasser, J. P., Blackburn, N., and van der Donk, W. A. (2003) SpaC and NisC, the cyclases involved in subtilin and nisin biosynthesis, are zinc proteins, *Biochemistry 42*, 13613-13624.

19. Li, B., Yu, J. P., Brunzelle, J. S., Moll, G. N., van der Donk, W. A., and Nair, S. K. (2006) Structure and mechanism of the lantibiotic cyclase involved in nisin biosynthesis, *Science 311*, 1464-1467.

20. Hightower, K. E., and Fierke, C. A. (1999) Zinc-catalyzed sulfur alkylation: insights

from protein farnesyltransferase, *Curr. Opin. Chem. Biol. 3*, 176-181.

21. Johnston, C. A., Temple, B. R., Chen, J. G., Gao, Y., Moriyama, E. N., Jones, A. M., Siderovski, D. P., and Willard, F. S. (2007) Comment on "A G protein coupled receptor is a plasma membrane receptor for the plant hormone abscisic acid", *Science 318*, 914; author reply 914.

22. Bauer, H., Mayer, H., Marchler-Bauer, A., Salzer, U., and Prohaska, R. (2000) Characterization of p40/GPR69A as a peripheral membrane protein related to the lantibiotic synthetase component C, *Biochem. Biophys. Res. Commun. 275*, 69-74.

23. Mayer, H., Bauer, H., Breuss, J., Ziegler, S., and Prohaska, R. (2001) Characterization of rat LANCL1, a novel member of the lanthionine synthetase C-like protein family, highly expressed in testis and brain, *Gene 269*, 73-80.

24. Chatterjee, C., Miller, L. M., Leung, Y. L., Xie, L., Yi, M., Kelleher, N. L., and van der Donk, W. A. (2005) Lacticin 481 synthetase phosphorylates its substrate during lantibiotic production, *J. Am. Chem. Soc. 127*, 15332-15333.

25. Kodani, S., Hudson, M. E., Durrant, M. C., Buttner, M. J., Nodwell, J. R., and Willey, J. M. (2004) The SapB morphogen is a lantibiotic-like peptide derived from the product of the developmental gene *ramS* in *Streptomyces coelicolor*, *Proc. Natl. Acad. Sci. U.S.A. 101*, 11448-11453.

26. Goto, Y., Li, B., Claesen, J., Shi, Y., Bibb, M. J., and van der Donk, W. A. (2010) Discovery of unique lanthionine synthetases reveals new mechanistic and evolutionary insights, *PLoS Biol. 8*, e1000339.

27. Müller, W. M., Schmiederer, T., Ensle, P., and Süssmuth, R. D. (2010) In vitro biosynthesis of the prepeptide of type-III lantibiotic labyrinthopeptin A2 including formation of a C-C bond as a post-translational modification, *Angew. Chem., Int. Ed. 49*, 2436-2440.

28. Goto, Y., Ökesli, A., and van der Donk, W. A. (2011) Mechanistic studies of Ser/Thr dehydration catalyzed by a member of the LanL lanthionine synthetase family, *Biochemistry 50*, 891-898.

29. Li, H., Xu, H., Zhou, Y., Zhang, J., Long, C., Li, S., Chen, S., Zhou, J. M., and Shao, F. (2007) The phosphothreonine lyase activity of a bacterial type III effector family, *Science 315*, 1000-1003.

30. Wang, H., and van der Donk, W. A. (2012) Biosynthesis of the Class III Lantipeptide Catenulipeptin, *ACS Chem. Biol. 7*, 1529-1535.

31. Krawczyk, B., Völler, G. H., Völler, J., Ensle, P., and Süssmuth, R. D. (2012) Curvopeptin: a new lanthionine-containing class III lantibiotic and its co-substrate promiscuous synthetase, *ChemBioChem 13*, 2065-2071.

32. Jenke-Kodama, H., Müller, R., and Dittmann, E. (2008) Evolutionary mechanisms underlying secondary metabolite diversity, *Prog. Drug. Res. 65*, 119, 121-140.

33. Toogood, P. L. (1993) Model Studies of Lantibiotic Biogenesis, *Tetrahedron Lett. 34*, 7833-7836.

34. Okeley, N. M., Zhu, Y., and van der Donk, W. A. (2000) Facile chemoselective synthesis of dehydroalanine-containing peptides, *Org. Lett. 2*, 3603-3606.

35. Burrage, S., Raynham, T., Williams, G., Essex, J. W., Allen, C., Cardno, M., Swali, V., and Bradley, M. (2000) Biomimetic synthesis of lantibiotics, *Chem.-Eur. J. 6*, 1455-1466.

36. Zhou, H., and van der Donk, W. A. (2002) Biomimetic stereoselective formation of methyllanthionine, *Org. Lett. 4*, 1335-1338.

37. Zhu, Y., Gieselman, M., Zhou, H., Averin, O., and van der Donk, W. A. (2003) Biomimetic studies on the mechanism of stereoselective lanthionine formation, *Org. Biomol. Chem. 1*, 3304-3315.

38. Zhang, Q., Yu, Y., Velásquez, J. E., and van der Donk, W. A. (2012) Evolution of lanthipeptide synthetases, *Proc. Natl. Acad. Sci. U. S. A. 109*, 18361-18366.

39. Yang, X., and van der Donk, W. A. (2013) Ribosomally synthesized and post-translationally modified peptide natural products: new insights into the role of leader and core peptides during biosynthesis, *Chemistry 19*, 7662-7677.

40. Lubelski, J., Khusainov, R., and Kuipers, O. P. (2009) Directionality and Coordination of Dehydration and Ring Formation during Biosynthesis of the Lantibiotic Nisin, *J. Biol. Chem. 284*, 25962-25972.

41. Lee, M. V., Ihnken, L. A., You, Y. O., McClerren, A. L., van der Donk, W. A., and Kelleher, N. L. (2009) Distributive and directional behavior of lantibiotic synthetases revealed by high-resolution tandem mass spectrometry, *J. Am. Chem. Soc. 131*, 12258-12264.

42. Thibodeaux, C. J., Ha, T. J., and van der Donk, W. A. (2014) A Price to Pay for Relaxed Substrate Specificity: A Comparative Kinetic Analysis of the Class II Lanthipeptide Synthetases ProcM and HalM2, *J. Am. Chem. Soc. 136*, 17513-17529.

43. Krawczyk, B., Ensle, P., Müller, W. M., and Süssmuth, R. D. (2012) Deuterium Labeled Peptides Give Insights into the Directionality of Class III Lantibiotic Synthetase LabKC, *J. Am. Chem. Soc. 134*, 9922.

44. van den Hooven, H. W., Rollema, H. S., Siezen, R. J., Hilbers, C. W., and Kuipers, O. P. (1997) Structural features of the final intermediate in the biosynthesis of the lantibiotic nisin. Influence of the leader peptide, *Biochemistry 36*, 14137-14145.

45. Beck-Sickinger, A. G., and Jung, G. (1991) Synthesis and conformational analysis of lantibiotic leader-, pro-, and pre-peptides, In *Nisin and Novel Lantibiotics* (Jung, G., and Sahl, H.-G., Eds.), pp 218-230, ESCOM, Leiden.

46. Kuipers, A., De Boef, E., Rink, R., Fekken, S., Kluskens, L. D., Driessen, A. J., Leenhouts, K., Kuipers, O. P., and Moll, G. N. (2004) NisT, the transporter of the lantibiotic nisin, can transport fully modified, dehydrated and unmodified prenisin and fusions of the leader peptide with non-lantibiotic peptides, *J. Biol. Chem. 279*, 22176-22182.

47. Rink, R., Wierenga, J., Kuipers, A., Kluskens, L. D., Driessen, A. J. M., Kuipers, O. P., and Moll, G. N. (2007) Production of dehydroamino acid-containing peptides by *Lactococcus lactis*, *Appl. Environ. Microbiol. 73*, 1792-1796.

48. Kluskens, L. D., Kuipers, A., Rink, R., de Boef, E., Fekken, S., Driessen, A. J., Kuipers, O. P., and Moll, G. N. (2005) Post-translational Modification of Therapeutic Peptides By NisB, the Dehydratase of the Lantibiotic Nisin, *Biochemistry 44*, 12827-12834.

49. Levengood, M. R., Patton, G. C., and van der Donk, W. A. (2007) The leader peptide is not required for post-translational modification by lacticin 481 synthetase, *J. Am. Chem. Soc. 129*, 10314-10315.

50. Oman, T. J., Knerr, P. J., Bindman, N. A., Velasquez, J. E., and van der Donk, W. A. (2012) An engineered lantibiotic synthetase that does not require a leader peptide on its substrate, *J. Am. Chem. Soc. 134*, 6952–6955.

51. Clatworthy, A. E., Pierson, E., and Hung, D. T. (2007) Targeting virulence: a new paradigm for antimicrobial therapy, *Nat. Chem. Biol. 3*, 541-548.

52. Chatterjee, C., Paul, M., Xie, L., and van der Donk, W. A. (2005) Biosynthesis and mode of action of lantibiotics, *Chem. Rev. 105*, 633-684.

53. Shivange, A. V., and Daugherty, P. S. (2015) De novo discovery of bioactive cyclic peptides using bacterial display and flow cytometry, *Methods Mol. Biol. 1248*, 139-153.

54. Foster, A. D., Ingram, J. D., Leitch, E. K., Lennard, K. R., Osher, E. L., and Tavassoli, A. (2015) Methods for the Creation of Cyclic Peptide Libraries for Use in Lead Discovery, *J. Biomol. Screen. 20*, 563-570.

55. Field, D., Connor, P. M., Cotter, P. D., Hill, C., and Ross, R. P. (2008) The generation of nisin variants with enhanced activity against specific gram-positive pathogens, *Mol. Microbiol. 69*, 218-230.

56. Shi, Y., Yang, X., Garg, N., and van der Donk, W. A. (2011) Production of lantipeptides in *Escherichia coli*, *J. Am. Chem. Soc. 133*, 2338-2341.

CHAPTER II: CHIMERIC PEPTIDES WITH PROCA LEADER PEPTIDE*

2.1 INTRODUCTION

Prochlorosins were discovered in 2010 from the planktonic marine cyanobacterium *Prochlorococcus* MIT9313 in a genome-mining effort to search for novel lanthipeptides in the unexplored genomes.[1] Remarkably, the genome of *Prochlorococcus* MIT9313 contains only one *lanM* gene encoding ProcM but 30 *lanA* genes encoding precursor peptides.[1, 2] The ProcA peptides have a high level of conservation in the N-terminal leader region and hypervariability in the C-terminal core region (Figure 2.1). The activity of ProcM has been confirmed for all 18 precursor peptides investigated thus far.[1-3] Giving the high degree of sequence conservation in the leader peptides, it is very likely that all 30 ProcAs are substrates of ProcM. In contrast, most LanM enzymes that have been investigated have only a single substrate.[4] The molecular details that allow this high substrate tolerance are largely unknown.

---

*Reproduced in part with permission from Yu, Y., Zhang, Q., and van der Donk, W. A. (2013) Insights into the evolution of lanthipeptide biosynthesis, *Protein Sci. 22*, 1478-1489, Zhang, Q., Yu, Y., Velásquez, J. E., and van der Donk, W. A. (2012) Evolution of lanthipeptide synthetases, *Proc. Natl. Acad. Sci. U. S. A. 109*, 18361-18366, and Yu, Y., Mukherjee, S., and van der Donk, W. A. (2015) Product Formation by the Promiscuous Lanthipeptide Synthetase ProcM is under Kinetic Control, *J. Am. Chem. Soc. 137*, 5140-5148.

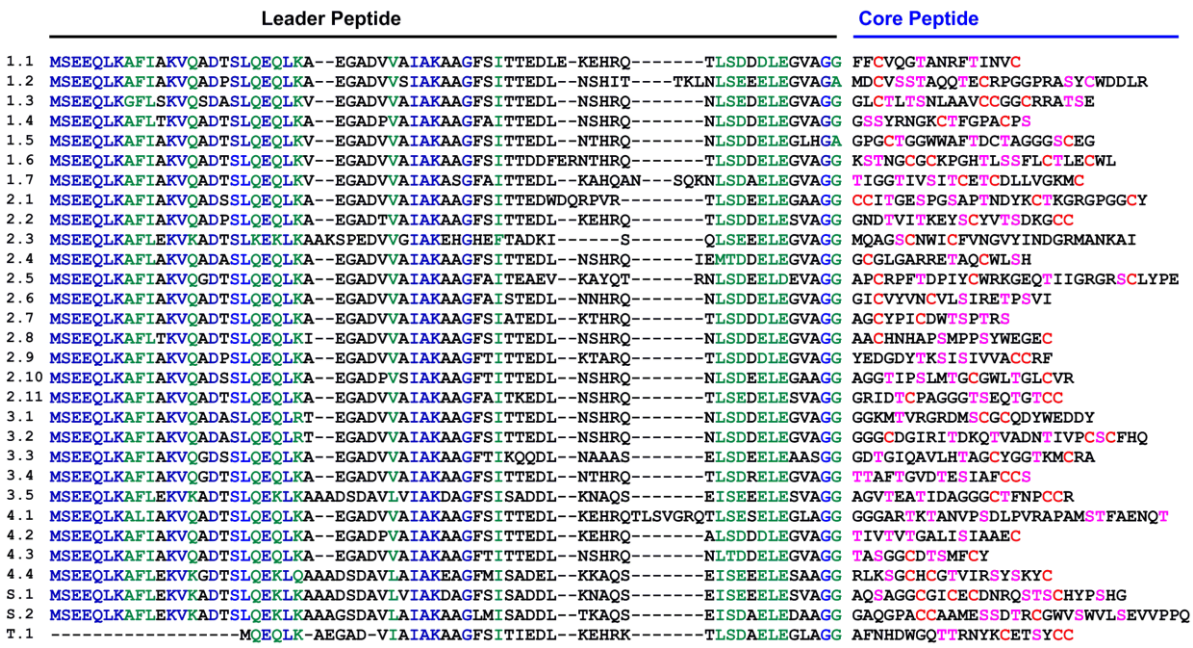

Figure 2.1. Sequence alignment of 30 ProcAs. The strictly conserved regions in the leader peptide are labeled in blue and other partially conserved regions in the leader peptide are labeled in green. Cysteines residues in the core peptide are labeled in red, serines and threonines are labeled in purple.

Discovery of this extraordinarily catalytic promiscuity of ProcM brought up the prospect of producing a variety of lanthipeptides utilizing the unique synthetase ProcM. Previous attempts in our lab to utilize the promiscuous catalytic activity of ProcM included constructing a fusion protein of ProcM with the ProcA leader peptide attached to it. This fusion protein was envisioned to catalyze the dehydration and cyclization of diverse peptides containing Ser, Thr, and Cys residues. This hypothesis was based on a previous study in our lab carried out by Dr. Oman, in which LctM fused with the LctA leader peptide catalyzed the modification of the LctA core peptide without its leader peptide attached to it.[5] But the fusion enzyme of ProcM did not exhibit similar activity. Thus, in this study, another approach was used, which involves

constructing chimeric peptides with a ProcA leader peptide attached to the target lanthipeptide core peptide. This investigation could have both scientific and commercial significance: First, producing lanthipeptides by co-expressing ProcM and chimeric peptides described above in *E. coli* instead of isolating these natural products from the producing strains is likely to greatly reduce the time and cost of the lanthipeptide production as well as increase the yield of the lanthipeptide;[6] Second, producing lanthipeptides using distinctly different lanthionine synthetases could provide insights about the roles of enzymes and substrates in the product formation processes.

In this chapter, efforts were made to use ProcM to synthesize nisin, a class I lanthipeptide of great commercial value,[7] epilancin 15X,[8] an effective antimicrobial peptide which also belongs to class I lanthipeptides, and lacticin 481,[9] a class II lanthipeptide originated from *Lactococcus lactis*. Nisin was the first member of the lanthipeptides to be characterized.[10] It is a highly effective antimicrobial peptide with a MIC value in the nanomolar range against a broad range of gram-positive bacteria. It has been widely used as a food preservative for half a century, but has not been reported to cause significant resistance in bacteria. The biosynthesis of nisin involves the dehydration of Ser/Thr residues catalyzed by the dehydratase NisB and formation of five thioether rings by the cyclase NisC (Figure 1.6 and Figure 2.2). Commercially, it is obtained by fermentation using its original producing strain *L. lactis*.[11] Epilancin 15X is a recently discovered lanthipeptide, which has impressive antibacterial activity against methicillin-resistant *Staphylococcus aureus* (MRSA) and vancomycin-resistant *enterococci* (VRE).[8] Besides the three thioether rings, it contains an unusual N-terminal *(R)*-2-hydroxypropionyl group (Lactate, Lac) (Figure 2.2). The Lac group potentially contributes to

the bioactivity of the lanthipeptide, but epilancin 15X without the N-terminal modification is also active against *L. lactis* HP.[8, 12, 13] Lacticin 481 (Figure 2.2) is a class II lanthipeptide produced by several strains of *L. lactis*, which has been suggested to display antimicrobial activities against gram-positive bacteria through the interaction with the membrane component lipid II.[14]

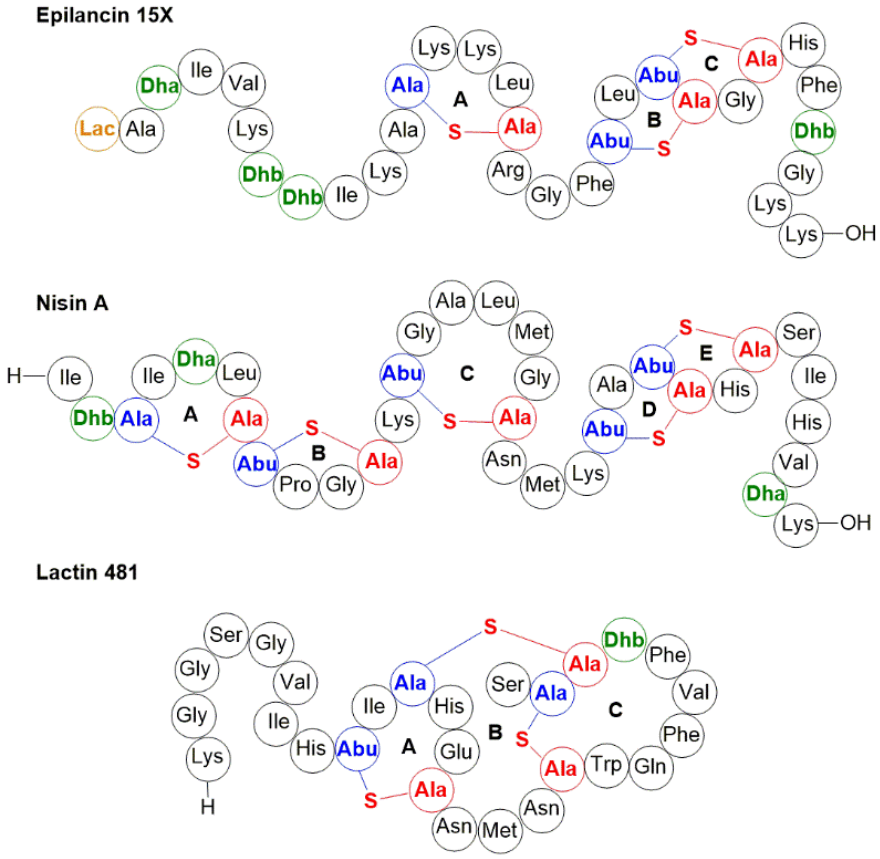

Figure 2.2. Structures of epilancin 15X, nisin and lacticin 481. Lan/MeLan segments derived from Cys are shown in red and originating from Ser/Thr are shown in blue. Additional posttranslational modifications are shown in orange. Dehydroamino acids are shown in green.

There have been extensive studies of utilizing the substrate-tolerance of lanthionine synthetases to produce lanthipeptide analogs or generate thioether crosslinks in

nonlanthipeptide targets.[15-18] Site-directed mutants of various lanthipeptide precursor peptides are recognized and successfully processed by the biosynthetic machinery in *in vitro* enzymatic assays as well as *in vivo* experiments with homologous and heterologous hosts, resulting in lanthipeptide variants.[15, 19-22] Two most extensive examples for the *in vivo* studies are complete alanine scanning of both peptides of the two-component lanthipeptide lacticin 3147[23] and the investigation of all possible single amino acid mutations for mersacidin.[24] Several of these *in vivo* studies have resulted in compounds with improved properties.[25, 26] *In vitro* studies have also demonstrated that core peptides containing nonproteinogenic amino acids can be modified to lantibiotic analogs with improved antimicrobial activities.[27] ProcM is the lanthionine synthetase identified thus far with the most diverse native substrate set. In this study, attempts were made to utilize ProcM for the production of both class I and class II lanthipeptides and expanding the possibility of a simplified route for lanthipeptide engineering.

2.2 RESULTS

2.2.1 The chimeric peptide strategy to utilize ProcM to synthesize other lanthipeptides

The leader peptides of the 30 ProcA precursor peptides are highly conserved, but the core peptides are hypervariable in terms of peptide length, amino acid composition, and positions of ring forming residues (Figure 2.1). Thus, it is reasonable to propose that the leader peptide is the crucial element for substrate-enzyme recognition while the core peptide is most likely to have little or no effect on the substrate-enzyme binding process. Chimeric peptides composed of one ProcA leader peptide (ProcA3.2 leader was used in this study) and the core peptide of the target lanthipeptides (NisA, ElxA and LctA core peptides were used in this study) were

constructed. It was envisioned that ProcM would likely accept these chimeric peptides as substrates and install thioether crosslinks in the core region. Engineering a hinge region incorporating recognition sites for endoproteases would enable the removal of the leader peptide after these modifications. If the native ring topologies of the target lanthipeptides are maintained by ProcM, bioactive products are proposed to be generated using this highly simplified and efficient method (Figure 2.3).

The construction of pET15b-*ProcALea_NisAStr*, pET15b-*ProcALea_ElxAStr* and pET15b-*ProcALea_LctAStr* was done by overlapping PCR (section 2.4.2). The *procM* gene was cloned into the MCS1 of pRSF-Duet1 vector to increase the yield of ProcM using procedures described in section 2.4.3. The peptide and protein overexpression and purification processes are described in section 2.4.5 and 2.4.6. *In vitro* enzymatic assays and digestion of the chimeric peptides were carried out as described in section 2.4.7. Samples were desalted using Zip-Tip, and analyzed by Matrix-Assisted Laser Desorption Ionization-Time-of-Flight (MALDI-TOF) mass spectrometry (MS). The cyclization of the peptides was evaluated using an iodoacetamide (IAA) assay described in section 2.4.8.

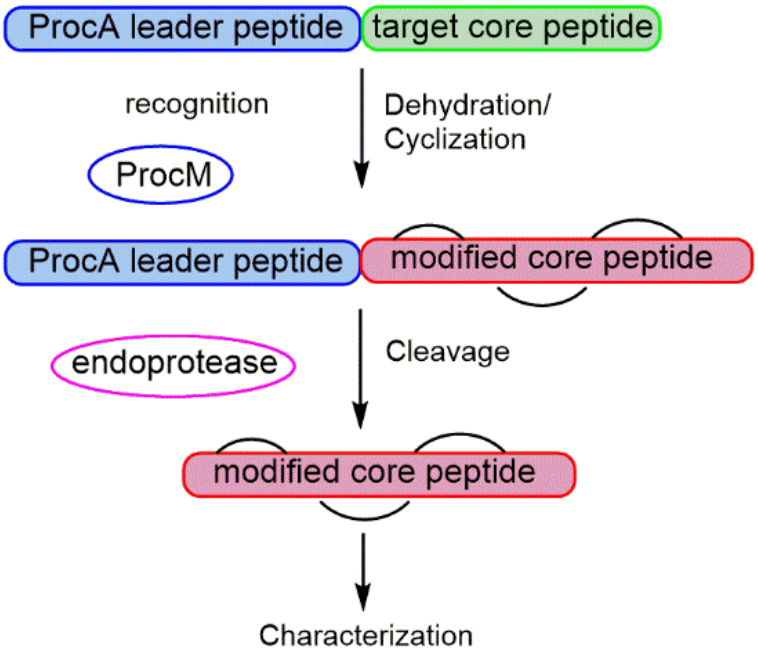

Figure 2.3. The chimeric peptide strategy utilizing ProcM to generate Lan/MeLan containing peptides.

2.2.2 Chimeric peptide with ProcA leader peptide and NisA core peptide

Incubation of the chimeric peptide composed of the ProcA3.2 leader peptide and NisA core peptide (ProcALea-NisAStr) with ProcM under the conditions listed in section 2.4.7 resulted in up to five-fold dehydration. During nisin biosynthesis, eight dehydrations and five cyclizations take place (Figure 1.6). To improve the extent of modifications, ProcALea_NisAStr peptide was also co-expressed with ProcM in *E. coli*. The *in vivo* co-expression resulted in up to 6-fold dehydration in the NisA core peptide (Figure 2.4). Using iodoacetamide (IAA) assay to probe for unreacted cysteines, it was shown that 3~4 rings were formed as the majority of the products reacted with either one or two molecules of IAA (Figure 2.5). As not all the five rings in nisin have known function in bioactivity, there was a possibility that the final product of the *in vivo* co-expression was still bioactive even though several Ser

or Thr residues were not dehydrated and one or two rings were not formed. In order to find out whether this situation was true, a protease cleavage site was introduced in the chimeric peptide to remove the leader peptide, which is known to keep the lantibiotics inactive.[28] But the mutated chimeric peptide ProcALea-NisAStr G−1K was not as good a substrate as the original chimeric peptide, with a series of one to five dehydrations observed under the same assay conditions. It is possible that the double-glycine motif might be important to guide the enzymatic modification.

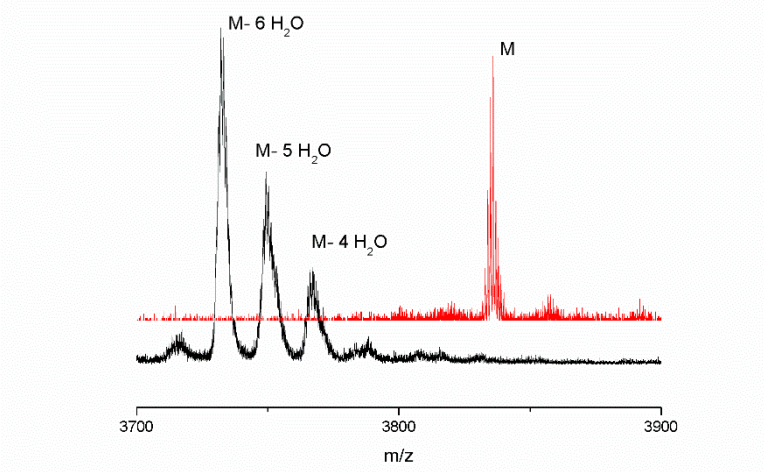

Figure 2.4. MALDI-TOF mass spectra of the ProcALea-NisAStr chimeric peptide modified by ProcM in *E. coli* and treated with GluC to remove most of the leader peptide (black line), and unmodified ProcALea-NisAStr chimeric peptide treated under the same conditions (red line). Calculated and observed masses are shown in Table 2.4.

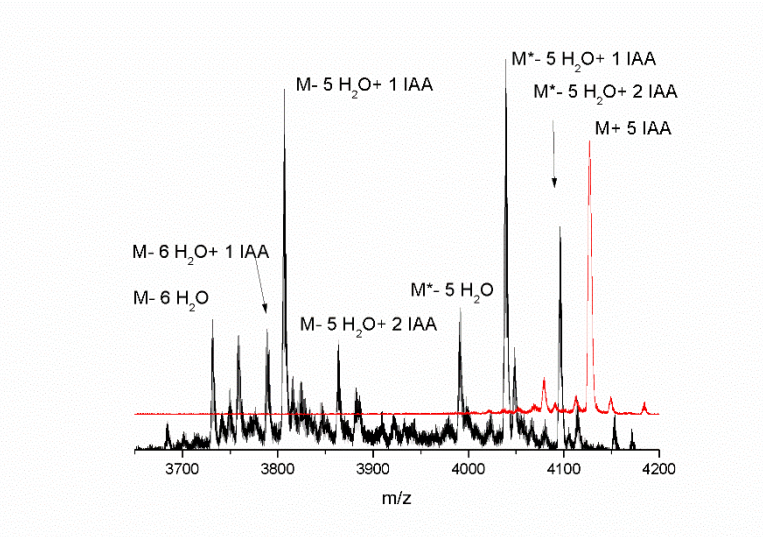

Figure 2.5. MALDI-TOF mass spectra of the ProcALea-NisAStr chimeric peptide modified by ProcM in *E. coli*, followed by subsequent digestion by GluC and treatment with IAA (black line), and unmodified ProcALea-NisAStr chimeric peptide treated under the same conditions (red line). Peaks labeled with the star symbol correspond to incomplete digestion when the glutamate at the -6 position in ProcA3.2 leader peptide skipped GluC digestion. Calculated and observed masses are shown in Table 2.4.

Therefore one more mutant, ProcALea-K-NisAStr, was made. In this mutant, a lysine residue was inserted after the double-Gly motif. ProcM also catalyzed up to 6 dehydrations and 3~4 cyclizations on this mutant *in vivo* (Figure 2.6 and Figure 2.7). The modified core peptide was protected from trypsin digestion, while the unmodified core peptide was not (Figure 2.8). The resulting core-containing peptides were tested for bioactivity, but no antimicrobial activity was observed against *L. lactis* HP even at 1.25 mM concentration (Figure 2.9).

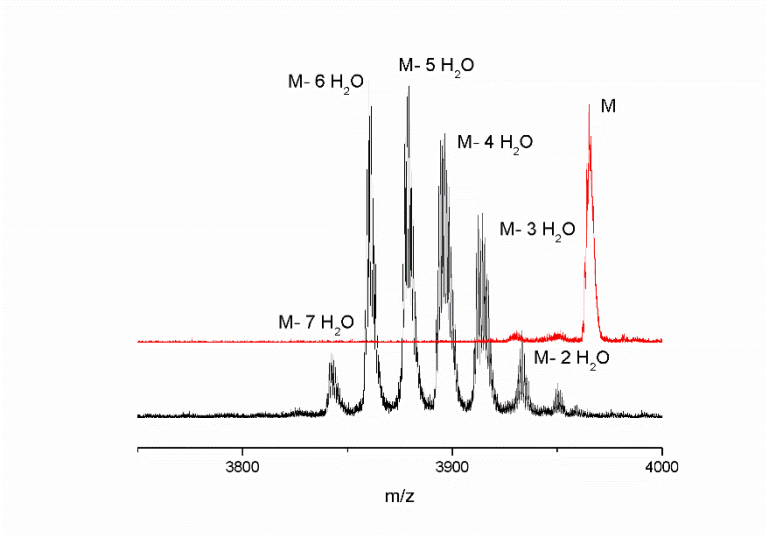

Figure 2.6. MALDI-TOF mass spectra of the ProcALea-K-NisAStr chimeric peptide modified by ProcM in *E. coli* and treated with GluC to remove most of the leader peptide (black line), and unmodified ProcALea-K-NisAStr chimeric peptide treated under the same conditions (red line). Calculated and observed masses are shown in Table 2.4.

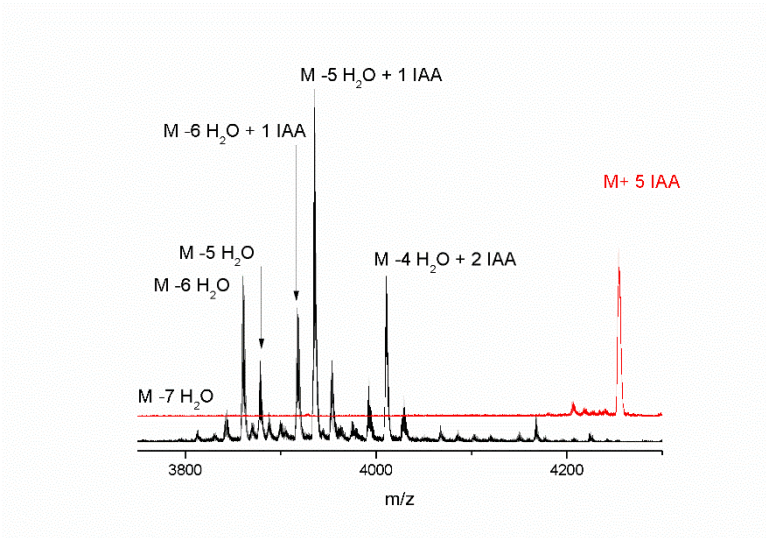

Figure 2.7. MALDI-TOF mass spectra of the ProcALea-K-NisAStr chimeric peptide modified by ProcM in *E. coli*, followed by subsequent digestion by GluC and treatment with IAA (black line), and unmodified ProcA Lea-K-NisAStr chimeric peptide treated under the same conditions (red line). Calculated and observed masses are shown in Table 2.4.

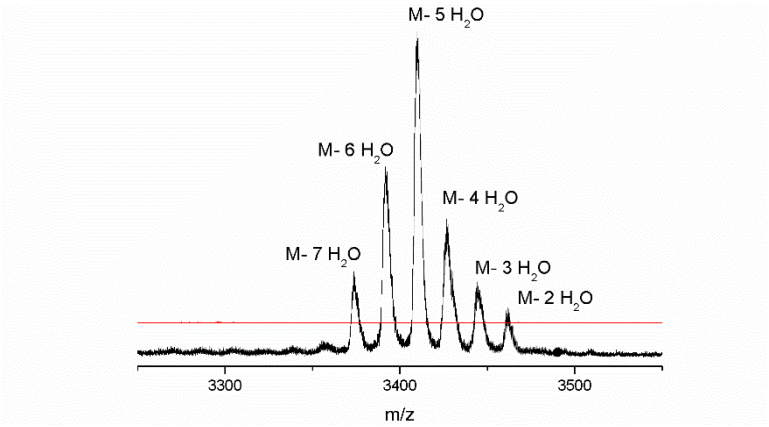

Figure 2.8. MALDI-TOF mass spectra of the ProcALea-K-NisAStr chimeric peptide modified by ProcM in *E. coli* and treated with trypsin to remove the leader peptide (black line), and unmodified ProcALea-K-NisAStr chimeric peptide treated under the same conditions (red line). Calculated and observed masses are shown in Table 2.4.

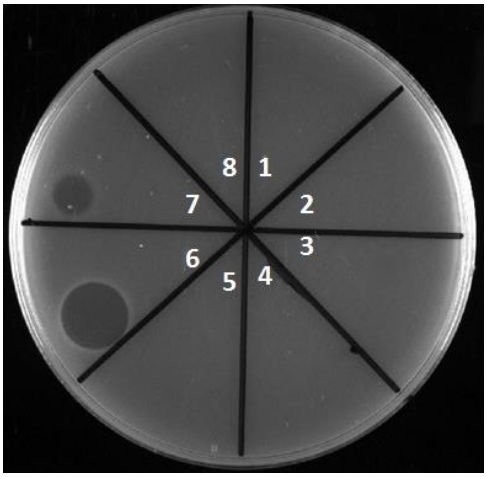

Figure 2.9. Antimicrobial assays against *L. lactis* HP. An aliquot of 5 μL of the following samples were spotted: (1) 125 μM ProcALea-K-NisAStr chimeric peptide modified by ProcM in *E. coli* and digested by trypsin; (2) 125 μM unmodified ProcALea-K-NisAStr chimeric peptide digested by trypsin; (3) 1.25 mM ProcALea-K-NisAStr chimeric peptide modified by ProcM in *E. coli* and digested by trypsin; (4) 1.25 mM unmodified ProcALea-K-NisAStr chimeric peptide digested by trypsin; (5) negative control with buffer for protease digestion; (6-8) 10, 1, and 0.1 μM authentic nisin.

2.2.3 Chimeric peptide with ProcA leader peptide and ElxA core peptide

As no mature nisin was produced using the chimeric peptide strategy, I looked at other candidates. The *in vitro* enzymatic assays of ProcM and its native substrates resulted in a maximum of five-fold dehydration in the ProcAs studied.[1] Analysis of the length of the core regions of the ProcAs and the numbers of maximum dehydrations that could occur suggested that most ProcM substrates were short and relatively simple. Therefore, we selected epilancin 15X, which is a lantibiotic of great potential application. It has only three non-overlapping rings although eight dehydrations occurred during its biosynthesis (Figure 2.2). However, both *in vivo* and *in vitro* modifications of ProcALea-E-ElxAStr (a glutamate residue was inserted between the ProcA3.2 leader peptide and ElxA core peptide) resulted in a similar MALDI-TOF mass spectrum, with the maximum number of dehydrations being five and the majority of the peptide dehydrated only four-fold (Figure 2.10). By subjecting the product to an IAA assay, it was shown that 2~3 rings were formed (Figure 2.11). However, the bioactivity assay using *L.lactis* HP as indicator strain resulted in no zone of growth inhibition even at 1.25 mM concentration, which suggested that epilancin 15X without the Lac group was not generated either due to incomplete cyclization or incorrect ring formation (Figure 2.12).

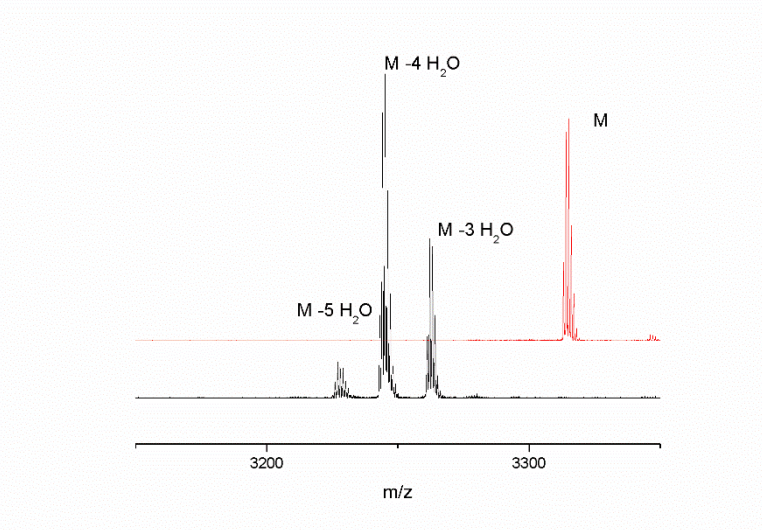

Figure 2.10. MALDI-TOF mass spectra of the ProcALea-E-ElxAStr chimeric peptide modified by ProcM in *E. coli* and treated with GluC to remove the leader peptide (black line), and unmodified ProcALea-E-ElxAStr chimeric peptide treated under the same conditions (red line). Calculated and observed masses are shown in Table 2.4.

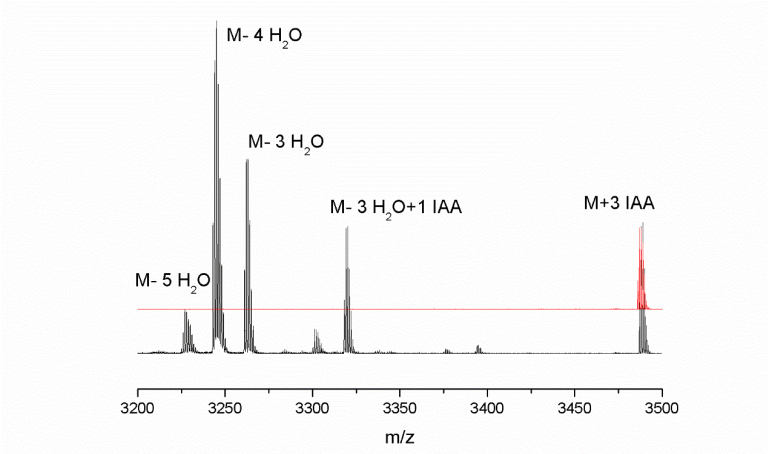

Figure 2.11. MALDI-TOF mass spectra of the ProcALea-E-ElxAStr chimeric peptide modified by ProcM in *E. coli*, followed by subsequent digestion by GluC and treatment with IAA (blank line), and unmodified ProcALea-E-ElxAStr chimeric peptide treated under the same conditions (red line). Calculated and observed masses are shown in Table 2.4.

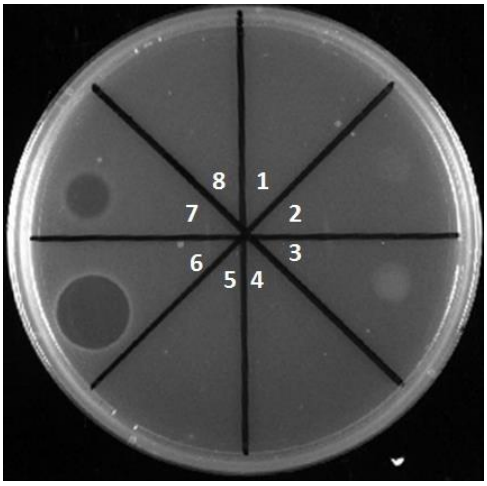

Figure 2.12. Antimicrobial assays against *L. lactis* HP. An aliquot of 5 μL of the following samples were spotted: (1) 125 μM ProcALea-E-ElxAStr chimeric peptide modified by ProcM in *E. coli* and digested by GluC; (2) 500 μM unmodified ProcALea-E-ElxAStr chimeric peptide digested by GluC; (3) 1.25 mM ProcALea-E-ElxAStr chimeric peptide modified by ProcM in *E. coli* and digested by GluC; (4) 1.25 mM unmodified ProcALea-E-ElxAStr chimeric peptide digested by trypsin; (5) negative control with buffer for protease digestion; (6-8) 10, 1, and 0.1 μM authentic nisin.

2.2.4 Chimeric peptide with ProcA leader peptide and LctA core peptide

The two class I lanthipeptide I selected in this study both require eight-fold dehydration in the core region to generate mature peptides, however, ProcM only showed a maximum of five-fold dehydration on its native substrates.[1] Therefore, the next candidate I studied was the class II lanthipeptide lacticin 481, which only requires four dehydrations and three cyclizations. A variety of other lantibiotics contain a similar ring topology as lacticin 481, which is often referred to as the lacticin 481 group.[29] The same strategy could easily be expanded for the production and engineering of these lanthipeptides.

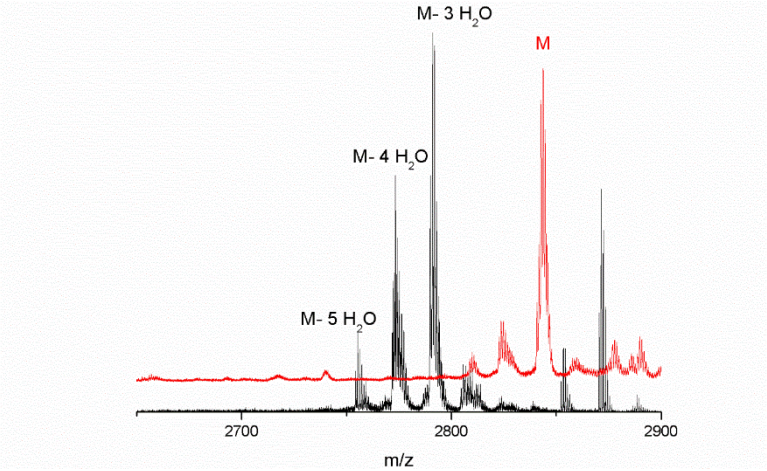

Figure 2.13. MALDI-TOF mass spectra of the ProcALea-LctAStr chimeric peptide modified by ProcM in *E. coli*, followed by subsequent digestion by LysC and treatment with IAA (black line), and unmodified ProcALea-LctAStr chimeric peptide treated under the same conditions (red line). Calculated and observed masses are shown in Table 2.4.

The *in vivo* modifications of ProcALea-LctAStr resulted in up to five dehydrations in the core peptide with the majority of the peptide dehydrated three-fold (Figure 2.13). As the wild-type lacticin 481 was dehydrated four times, it is likely that Ser4 residue, which usually skips dehydration, was also modified. IAA assay demonstrated that 2~3 rings were formed (Figure 2.14). The tandem MS data suggested for some of the products, the large overlapping ring was not formed (Figure 2.15). Interestingly, the bioactivity assay using the *in vivo* modified ProcALea-LctAStr digested with trypsin produced a clear zone of inhibition, which suggests the generation of wild-type lacticin 481 (Figure 2.16). On the other hand, the *in vitro* modification of ProcALea-LctAStr did not result in any bioactive product, which is consistent with the observation of higher modification efficiency in the in *vivo* assay than *in vitro* experiments.

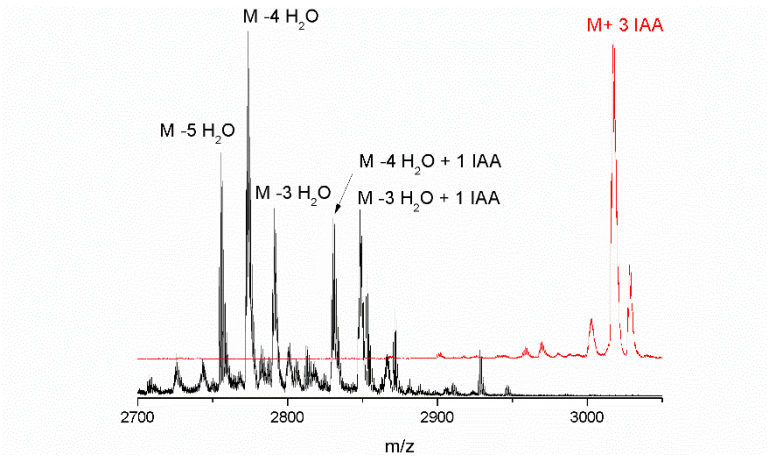

Figure 2.14. MALDI-TOF mass spectra of the ProcALea-LctAStr chimeric peptide modified by ProcM in *E. coli*, followed by subsequent digestion by LysC and treatment with IAA (black line), and unmodified ProcALea-LctAStr chimeric peptide treated under the same conditions (red line). Calculated and observed masses are shown in Table 2.4.

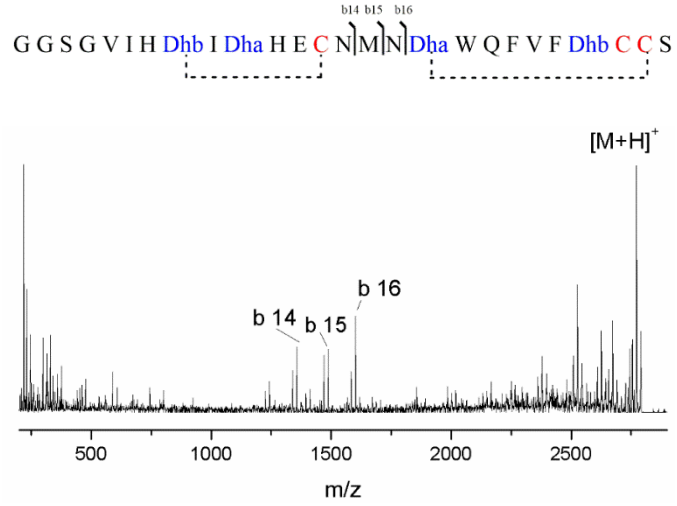

Figure 2.15. ESI-MSMS of the 4-fold dehydrated species of the ProcALea-LctAStr chimeric peptide modified by ProcM in *E. coli*, digested by trypsin and purified with analytical HPLC using a C18 column. The residues connected with dashed lines are potentially connected by a thioether crosslink.

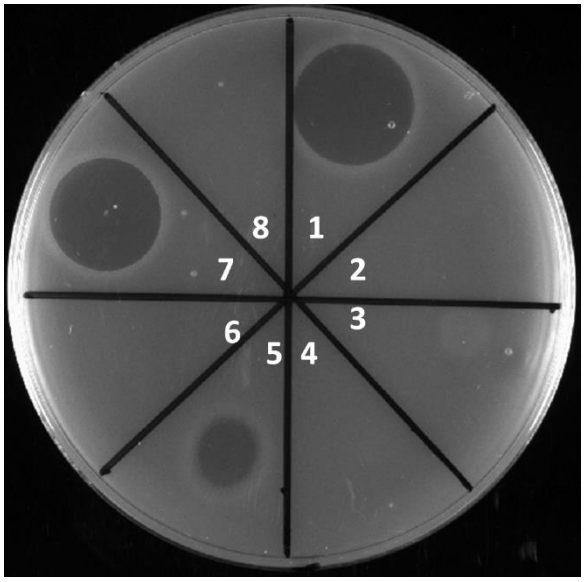

Figure 2.16. Antimicrobial assays against *L. lactis* HP. An aliquot of 5 μL of the following samples were spotted: (1) 250 μM ProcALea-LctAStr chimeric peptide modified by ProcM in *E. coli* and digested by trypsin ; (2) 250 μM unmodified ProcALea-LctAStr chimeric peptide digested by trypsin; (3) 250 μM ProcALea-LctAStr chimeric peptide modified by ProcM *in vitro* and digested by trypsin; (4) *in vitro* enzymatic assay mixture without ProcM enzyme and digested by trypsin.; (5) 10 μM authentic lacticin 481; (6) blank; (7) 10 μM LctA modified by LctM *in vitro* and digested by trypsin; (8) blank. Treatment with trypsin generated Δ1-lacticin lacking the N-terminal Lys residue. Previous studies have shown that Δ1-lacticin is about 3-fold less active than authentic lacticin 481.[27]

2.3 DISCUSSION AND OUTLOOK

In this study, it is demonstrated that the highly promiscuous class II lanthionine synthetase ProcM was able to install thioether crosslinks in the core region of both class I and class II lanthipeptides that are attached to a ProcA leader peptide. In the case of two relatively long class I lanthipeptides, which require eight dehydrations for maturation, ProcM showed limited

capability to modify these core peptides to full extent. On the other hand, in the case of ProcALea-LctAStr, the *in vivo* modifications of ProcM in *E. coli* resulted in the successful generation of bioactive lacticin 481, which could be potentially applied to the production of relatively short lanthipeptides with a small number of rings. Moreover, this study provides insights on the evolutionary story of lanthionine synthetases and substrates.

The availability of the rapidly growing number of bacterial genomes is starting to provide the first insights into the possible mechanisms by which lanthipeptide biosynthesis may have evolved. Most of the "modern" lanthipeptide synthetases are highly evolved and make single polycyclic products that are the result of a remarkably well-orchestrated set of reactions. The discovery of class III and IV lanthipeptide synthetases (chapter I) suggests that the biosynthetic machinery may have evolved from stand-alone kinases, phosphoThr/phosphoSer lyases, and Zn^{2+}-dependent enzymes with a very common-toroidal core structure. The low sequence specificity of primitive progenitor enzymes appears to have been maintained with respect to the phosphorylation and elimination reactions, but an allosteric regulation mechanism that depends on a leader peptide has evolved that assures that the dehydratases do not non-specifically act on any nascent polypeptide produced by the ribosome. The evolutionary origin of class I and II dehydratase machinery is less clear, but both involve activation of the Ser and Thr side chains, and presumably they have evolved from other enzymes that utilized nucleotides and Glu-tRNA. LanM enzymes from various phyla that like ProcM have three cysteines ligated to the active site zinc ion (as opposed to the more common Cys-Cys-His ligand set) cluster together in a phylogenetic tree.[4, 13] Most importantly, the phylogenic data suggest that for some scaffolds, the ring topology of the final lanthipeptides may be determined

in part by the sequence of the precursor peptides and not just by the biosynthetic enzymes. This notion was supported by the studies in this chapter with chimeric peptides, suggesting that prochlorosin biosynthetic enzymes can produce the correct ring topology of lacticin 481. These results highlight the potential of lanthipeptide synthetases for bioengineering and combinatorial biosynthesis.

The evolution of the precursor peptides and hence the structure of the final products is complex. For some systems, the progenitors of the precursor peptides can be inferred from the sequences of their leader peptides, with a group of class II peptides in cyanobacteria having descended from nitrogen-fixing proteins.[30] But for the majority of lanthipeptide precursor peptides, their origins are unclear. Intriguing potential insights into the evolution of the posttranslational modification process that turns a linear peptide into a highly structured product that recognizes its target with exquisite selectivity[31, 32] is provided by a subgroup of class II lanthipeptides in cyanobacteria that includes the prochlorosins.[1] The observation of a large number of precursor peptides with highly conserved leader peptides but highly diverse core peptides that are converted into a library of products by one synthetase suggests that this may be a snap-shot of a system under evolution featuring a non-specific synthetase that can act on a large number of core peptide sequences. Some support has been obtained that the topology and stereochemistry generated by such a synthetase may be determined at least in part by inherent conformational preferences of the core peptide sequence, perhaps reinforced by interaction of the core peptide with the molecular surface of the synthetase. During evolution, such ancestral synthetases would have evolved to more specific enzymes when they made products that were beneficial to the organism, ultimately resulting in proteins that make one

product very efficiently. Many questions still remain for this model including how a non-specific synthetase is maintained, whether such a synthetase is primitive or has evolved to have low substrate specificity, and how the biosynthetic enzymes interact with their leader and core peptides. Future studies will hopefully provide answers to these questions. The intriguing questions as to how these enzymes can tolerate so many different substrates is addressed in chapter III.

2.4 EXPERIMENTAL PROCEDURES

2.4.1 Materials

All oligonucleotides were purchased from Integrated DNA Technologies and used as received. Restriction endonucleases, DNA polymerases and T4 DNA ligase were from New England Biolabs. Media components were obtained from Difco laboratories. *E. coli* DH5α cells were used as host for cloning and plasmid propagation, and *E. coli* BL21 (DE3) or *E. coli* Rosetta 2 (DE3) cells were used as host for expression. Endoproteases were ordered from Roche Biosciences or New England Biolabs. Unless specified otherwise, chemicals were purchased from Sigma Aldrich or Fisher Scientific.

2.4.2 Construction of the expression plasmids for chimeric peptides

The plasmids pET15b-*procA3.2*, pET15b-*procA*3.3[1] pET15b-*lctA*,[4] and pET15b-*nisA*[33] were isolated from *E. coli* DH5α cells and used as templates for PCR reactions. The *elxA* gene was amplified from the genomic DNA of the producing strain *Staphylococcus epidermidis* 15X154. The DNA sequence for each chimeric peptide was amplified using overlapping PCR. In the first round PCR, the DNA sequence for leader peptide and core peptide were amplified

individually using one end primer containing the restriction site and one connection primer containing the C-terminal sequence of the leader and the N-terminal sequence of core peptide (Table 2.1). The PCR reactions were carried out using Phusion polymerase and 30 cycles of denaturing (94 °C for 30 s), annealing (55 °C for 30 s), and extending (72 °C, 1 min for each 1.5 kb). The second round PCR was carried out using DNA purified from the first round PCR without any primers for 7 cycles to anneal the leader and core fragments, and with two end primers containing restriction sites for an additional 23 cycles to amplify the gene for the chimer peptide. The final PCR products were digested with NdeI and XhoI restriction enzymes and ligated into the pET15b vector to generate the constructs listed in Table 2.1. The sequences of the inserts were checked by DNA sequencing.

Table 2.1. Primers used in section 2.4.2.

Construct	**Primer Name**	**Primer Sequences (5'-3')**
pET15b-*procALea-nisAStr* (Template: pET15b-*procA3.2*, pET15b-*nisA*)	Forward-1: *procA3.2*_NdeI_FP	TCAGATCATATGATGTCAGAAG AACAACTC
	Reverse-1: *procA3.2Lea_nisAStr*_ Conn_RP	CATAGCGAAATACTTGTAATTC CCCCAGCCACACCTTC
	Forward-2: *procA3.2Lea_nisAStr* _Conn_FP	GAAGGTGTGGCTGGGGGAATT ACAAGTATTTCGCTATG
	Reverse-2: *nisA*_XhoI_RP	ACAGACGACTCGAGTTATTTGC TTACGTGAATACTACAATGACA AG
pET15b-*procALea-K-nisAStr* (Template: pET15b-*procA3.2*, pET15b-*nisA*)	Forward-1: *procA3.2*_NdeI_FP	TCAGATCATATGATGTCAGAAG AACAACTC
	Reverse-1: *procA3.2Lea_K_ nisAStr*_Conn_RP	CGAAATACTTGTAATTTTTCCC CCAGCCACACCTTC
	Forward-2: *procA3.2Lea_K_ nisAStr* _Conn_FP	GAAGGTGTGGCTGGGGGAAAA ATTACAAGTATTTCG
	Reverse-2: *nisA*_XhoI_RP	ACAGACGACTCGAGTTATTTGC TTACGTGAATACTACAATGACA AG
pET15b-*procALea-E-ElxA* (Template: pET15b-*procA3.2*, genomic DNA from *Staphylococcus epidermidis* 15X154)	Forward-1: *procA3.2*_NdeI_FP	TCAGATCATATGATGTCAGAAG AACAACTC
	Reverse-1: *procA3.2Lea_E_ ElxAStr*_Conn_RP	AACAATACTAGCTGATTCTCCC CCAGCCACACC
	Forward-2: *procA3.2Lea_E_ ElxAStr*_Conn_FP	GGTGTGGCTGGGGGAGAATCA GCTAGTATTGTT
	Reverse-2: *ElxA*_XhoI_RP	GCCTCGAGTTATTTTTTACCAG TAAAGTG

(Table 2.1 continued)

pET15b-*procALea-lctAStr* (Template: pET15b-*procA3.2*, pET15b-*lctA*)	Forward-1: *procA3.2*_NdeI_FP	TCAGATCATATGATGTCAGAAG AACAACTC
	Reverse-1: *procA3.2Lea_lctAStr* _Conn_RP	CTCCACTGCCGCCTTTTCCCCC AGCCACACC
	Forward-2: *procA3.2Lea_lctAStr* _Conn_FP	GGTGTGGCTGGGGGAAAAGGC GGCAGTGGAG
	Reverse-2: *lctA*_XhoI_RP	TCAGATCTCGAGTTAAGAGCA GCAAGTAAATAC

2.4.3 Construction of pRSF-Duet1-*procM*

The plasmid pET28b-*procM* was isolated from *E. coli* DH5α cells.[1] The *procM* gene was PCR amplified using Phusion polymerase by 30 cycles of denaturing (94 °C for 30 s), annealing (55 °C for 30 s), and extending (72 °C for 2 min) using *procM*_Duet1_EcoRI_FP as forward primer, *procM*_Duet1_NotI_RP as reverse primer (Table 2.2), and pET28b-*procM* as template. The PCR product was digested with EcoRI and NotI restriction enzymes and ligated into the first multiple cloning site (MCS1) in the pRSF-Duet1 vector to generate the construct pRSF-Duet1-*procM* (MCS1). The sequence of the *procM* insert was checked by DNA sequencing.

Table 2.2. Primers used in section 2.4.3.

Primer Name	**Primer Sequences (5'-3')**
procM _EcoRI_FP	CAGGATCCGAATTCGATGGAAAGTCCATCATCTTG G
procM _NotI_RP	AAGGAAAAAAGCGGCCGCTTATTCAGTAGGCCAG AGA

2.4.4 Construction of the co-expression plasmids for chimeric peptides and ProcM

The *procM* gene was amplified by PCR in 30 cycles of denaturing (98 °C for 30 s), annealing (55 °C for 30 s), and extending (72 °C, 1 min for each 1.5 kb) using the plasmid pRSF-Duet1-*procM* as templates and using *procM*_Duet1_NdeI_FP as forward primer and *procM*_Duet1_KpnI_RP as reverse primer (Table 2.3). The PCR products were digested with NdeI and KpnI restriction enzymes and ligated into the second multiple cloning site (MCS2) in the pRSF-Duet1 vector to generate the construct pRSF-Duet1-*procM* (MCS2). The sequences of the resulting products were confirmed by DNA sequencing.

Chimeric genes were amplified by PCR using 30 cycles of denaturing (98 °C for 10 s), annealing (55 °C for 30 s), and extending (72 °C for 15 s) using the primers listed in Table 2.3 and the constructs generated in section 2.4.2 as templates. The resulting DNA inserts and the pRSF-Duet1 vector containing *procM* in MCS2 were digested with EcoRI and NotI and ligated. The sequences of the resulting products were confirmed by DNA sequencing.

Table 2.3. Primers used in section 2.4.4.

Primer Name	**Primer sequence (5'-3')**
*procM*_Duet1_NdeI_FP	GGTTGGTTCATATGGAAAGTCCATCATCTTGG
*procM*_Duet1_KpnI_RP	AAGTAGTTGGTACCTTATTCAGTAGGCCAGAGAC
*procA*3.2_EcoRI_FP	CAGGATCCGAATTCGATGTCAGAAGAACAACTCAAGG
*nisA*_NotI_RP	AAGGAAAAAAGCGGCCGCTTATTTGCTTACGTGAAT
*elxA*_NotI_RP	AAGGAAAAAAGCGGCCGCTTATTTTTTACCAGTAAAGT
*lctA*_NotI_RP	GGAAAAAAGCGGCCGCTTAAGAGCAGCAAGTA

2.4.5 Expression and purification of unmodified and modified peptides

E. coli BL21 (DE3) cells were transformed with pRSF-Duet1 constructs containing peptide and enzyme genes or pET15b constructs containing only the peptide gene via electroporation.

A single colony transformant was used to inoculate 20 mL of LB medium supplemented with 50 μg/mL kanamycin (pRSF-Duet1 constructs) or 100 μg/mL amplicilin (pET15b constructs). The culture was grown at 37 °C for 12 h and was used to inoculate 1 L of LB medium containing 50 μg/mL kanamycin or 100 μg/mL ampicillin and the cells were grown at 37 °C to $OD_{600} \approx$ 0.6. For the pRSF-Duet1 constructs, the culture was incubated at 4 °C on ice for 20 min, then IPTG was added to a final concentration of 0.5 mM, and the culture was incubated at 18 °C with shaking for an additional 16-20 h. For the pET15b constructs, IPTG was added to a final concentration of 1 mM and the culture was incubated at 37 °C with shaking for an additional 3 h. Cells were harvested by centrifugation at 12,000×*g* for 15 min at 4 °C, and the pellet was resuspended in 30 mL of LanA buffer A containing 6 M guanidine hydrochloride, 20 mM NaH_2PO_4, pH 7.5, 500 mM NaCl, and 0.5 mM imidazole. Cells were then lysed by sonication (5 min, 35% amplitude, 4.0 s ON, 9.9 s OFF) and the soluble and insoluble layers were separated by centrifugation at 23,700×*g* for 30 min at 4 °C. The unmodified peptides produced from pET15b constructs were found in the insoluble fraction, and the modified peptides produced by the pRSF-Duet1 co-expression constructs were mostly found in the soluble fraction. For the purification of unmodified peptides, the pellets were treated once more with LanA buffer A to remove any soluble proteins from the pellet. The pellet from the second wash was resuspended in 20 mL of LanA buffer B (6 M guanidine hydrochloride, 20 mM NaH_2PO4, 500 mM NaCl, 0.5 mM imidazole, pH 7.5). The insoluble portion was removed by centrifugation and the supernatant was clarified through 0.45 μm syringe filters prior to nickel affinity chromatography. For purification of modified peptides, the supernatant from the first sonication was directly filtered with 0.45 μm syringe filters. The resulting peptide solutions

were loaded onto a HisTrap nickel affinity column (GE healthcare), then the resin was washed with 2 column volumes (CV) each of LanA Buffer B and LanA Buffer C (4 M guanidine hydrochloride, 20 mM NaH_2PO_4, 300 mM NaCl, 30 mM imidazole, pH 7.5). The desired LanA peptide was eluted using 1–3 CV of LanA Buffer D (4 M guanidine hydrochloride, 20 mM Tris, 100 mM NaCl, 1 M imidazole, pH 7.5). Imidazole removal from the His_6-LanA peptides was achieved by preparative RP-HPLC using a Waters Delta-pak C4 column. Collected fractions were lyophilized and analyzed by MALDI-TOF MS.

2.4.6 Expression and purification of LanM enzymes

E. coli BL21 (DE3) or *E. coli* Rosetta 2 (DE3) cells were transformed via electroporation with pRSF-Duet1 or pET28b constructs containing the enzyme genes. A single colony transformant was used to inoculate a 50 mL of LB medium supplemented with 50 μg/mL kanamycin. The culture was grown at 37 °C for 12 h and was used to inoculate 4 L of LB medium containing 50 μg/mL kanamycin. Cells were grown at 37 °C to $OD_{600}\approx 0.6$. The culture was incubated on ice for 20 min, then IPTG was added to a final concentration of 0.5 mM, and the culture was incubated at 18 °C for an additional 16-20 h. Cells were harvested by centrifugation at 12,000×*g* for 15 min at 4 °C, and the pellets were frozen in liquid nitrogen and stored at −80 °C. All protein purification steps were performed at 4 °C. The cell pellets were suspended in LanM buffer A (20 mM Tris, 1 M NaCl, pH 8.0), and the cells were lysed using a high pressure homogenizer (Avestin, Inc.). Cell debris was pelleted via centrifugation at 23,700×*g* for 40 min at 4 °C. The supernatant was filtered using 0.45 μm syringe filters and injected onto an ÄKTA fast protein liquid chromatography (FPLC) system (GE Healthcare Life Sciences) equipped with a 5 mL HisTrap HP IMAC column previously charged with Ni^{2+} and equilibrated in start

buffer. The column was washed with 50 mL of start buffer and the protein was eluted using a gradient of 0-100% LanM buffer B (200 mM imidazole, 20 mM Tris, pH 8.0, 1 M NaCl) at a flow rate of 1.5 mL/min. UV absorbance at 280 nm was monitored and fractions were collected and analyzed by SDS-PAGE (4-20% Tris-glycine READY gel, BioRAD). The fractions containing ProcM or ProcM mutant proteins were combined and concentrated using an Amicon Ultra-15 Centrifugal Filter Unit (30, 50 or 100 kDa MWCO, Millipore). The concentrated protein sample was injected onto an ÄKTA FPLC system equipped with an XK16 16/60 column (GE Healthcare Life Sciences) packed with SuperDex 200 resin previously equilibrated in buffer A. The protein was eluted at a flow rate of 1 mL/min. UV absorbance at 280 nm was monitored and fractions were collected. Misfolded or aggregated protein was efficiently separated from soluble, correctly folded protein. The desired fractions were combined and concentrated using an Amicon Ultra-15 Centrifugal Filter Unit. Protein concentration was determined using a Bradford Assay Kit (Pierce) and $A280_{nm}$. Typical yields were 5~20 mg per liter of cell culture. The resulting samples were frozen and stored at −80 °C.

2.4.7 *In vitro* enzymatic assay and endoprotease digestion

The lyophilized peptides were dissolved in ddH_2O, and added to 100 μL of enzymatic assay solution. The final peptide and enzyme concentration varied because of different peptide solubility and reaction rates. The *in vitro* enzymatic assay included: 10 mM ATP, 1 mM TCEP, 50 mM HEPES (pH 8.0), 10 mM $MgCl_2$ and various amounts of peptide and enzyme. The reactions were incubated at room temperature for the time indicated in the figure legends and then quenched using 1% formic acid (final concentration). The pH was then adjusted to the optimal pH for the indicated endoprotease using 2.5 M NaOH. An aliquot of 2 μL of

endoprotease (LysC 6 U/μL, GluC 1 mg/mL, trypsin 1 mg/mL, AspN 1 mg/mL) was added to the solution and incubated for 1~3 h at room temperature or 37 °C. The resulting mixtures were desalted by ZipTip (Millipore) and analyzed by MALDI-TOF MS using an UltrafleXtreme instrument (Bruker).

2.4.8 Bioactivity assay

The growth inhibition activity of peptide sample was determined against *Lactococcus lactis HP.* A 25 mL aliquot of liquid molten GM17 agar (4% M17, 0.5% glucose, 1.5% agar) was cooled to 42 °C and seeded with 200 μL of overnight culture (OD_{600}=1) of the indicator strain *L. lactis HP*. After agar solidification in a Petri dish, peptide samples were added to the plate and dried near open flame for 1 hour. Plates were incubated for 15 h at 30°C and antibacterial activity was qualitatively determined by the presence or absence of a zone of growth inhibition.

2.4.9 Iodoacetamide assay (IAA assay)

An aliquot of 20 μL of the enzymatic reaction solution or untreated control peptide was incubated with 2 μL of 500 mM IAA, 1 μL of 50 mM TCEP, and 27 μL of 100 mM Tris-HCl (pH 8.9) at room temperature for 3 h. Protease was added to the mix and the sample was incubated for an additional 1~3 h in the dark at room temperature.

2.4.10 ESI–LC-MS and tandem MS

ESI-LC-MS and tandem MS analysis were performed with a Synapt ESI quadrupole TOF Mass Spectrometry System (Waters) coupled with an Acquity Ultra Performance Liquid Chromatography (UPLC) system (Waters). A sample of 5 μL of digested peptide was injected on a Phenomenex Jupiter C18 column, and separated by HPLC using a gradient of 2% mobile phase B to 40% mobile phase B over 20 min (mobile phase A: water with 0.1% formic acid,

mobile phase B: acetonitrile with 0.1% formic acid). Mass spectra were acquired in ESI positive mode in the range of 50-2000 m/z. The capillary voltage was 3500 V, and cone voltage was 40 V. The following parameters were also used: 120 ºC source temperature; 300 ºC desolvation temperature, 150 L/h cone gas flow, and 600 L/h desolvation gas flow. A transfer collision energy of 4 V was used for both MS and tandem MS, whereas the trap collision energy was set to 6 V for MS and a 20-50 V ramp for MSMS depending on the peptide. Glu-fibrinopeptide B (Sigma) was directly infused as lock mass. The tandem mass spectra were processed with MaxEnt3 and analyzed by Protein/Peptide Editor in BioLynx 4.1. The software for analyzing both precursor-ion and fragment-ion mass was set to report any mass within 0.3 amu of the calculated expected values.

2.4.11 Calculated and observed masses in the mass spectra

Table 2.4. Calculated and observed masses in the mass spectra.

***** monoisotopic masses plus one proton were shown for all peaks

Figure	Peak	Calculated mass	Observed mass
Figure 2.4 (Blank)	M - 6 H_2O	3729.7	3730.3
	M - 5 H_2O	3747.8	3747.9
	M - 4 H_2O	3765.8	3764.7
Figure 2.4 (Red)	M	3837.8	3833.9 (Disulfide bond formation between the cysteines)
Figure 2.5 (Blank)	M - 6 H_2O	3729.7	3729.2
	M - 6 H_2O + 1 IAA	3786.8	3785.8
	M - 5 H_2O + 1 IAA	3804.8	3804.2
	M - 5 H_2O + 2 IAA	3861.8	3861.0
Figure 2.5 (Red)	M + 5 IAA	4122.9	4124.9

(Table 2.4 continued)

Figure	Peak	Calculated mass	Observed mass
Figure 2.6 (Black)	M - 7 H_2O	3839.8	3840.3
	M - 6 H_2O	3857.8	3858.3
	M - 5 H_2O	3875.8	3876.4
	M - 4 H_2O	3893.9	3893.4
	M - 3 H_2O	3911.9	3911.0
Figure 2.6 (Red)	M	3965.9	3963.2 (Disulfide bond formation between the cysteines)
Figure 2.7 (Black)	M - 6 H_2O	3857.8	3858.4
	M - 5 H_2O	3875.8	3876.4
	M - 6 H_2O + 1 IAA	3914.9	3915.5
	M - 5 H_2O + 1 IAA	3932.9	3933.5
	M - 4 H_2O + 2 IAA	4007.9	4008.5
Figure 2.7 (Red)	M + 5 IAA	4251.01	4251.0
Figure 2.8 (Black)	M - 7 H2O	3370.6	3371.5
	M - 6 H2O	3388.6	3388.6
	M - 5 H2O	3406.6	3407.6
	M - 4 H2O	3424.6	3425.7
	M - 3 H2O	3442.6	3443.0
Figure 2.10 (Black)	M - 5 H2O	3224.7	3226.0
Figure 2.10 (Red)	M - 4 H2O	3242.7	3243.0
	M - 3 H2O	3260.8	3261.1
	M	3314.8	3313.1
Figure 2.11 (Black)	M - 5 H2O	3224.7	3226.0
	M - 4 H2O	3242.7	3243.0
	M - 3 H2O	3260.8	3261.0
	M - 3 H2O + 1 IAA	3317.8	3318.1
Figure 2.11 (Red)	M+ 3 IAA	3485.9	3486.0

(Table 2.4 continued)

Figure	Peak	Calculated mass	Observed mass
Figure 2.13 (Blank)	M - 5 H_2O	2754.1	2754.1
	M - 4 H_2O	2772.2	2772.2
	M - 3 H_2O	2790.2	2790.1
Figure 2.13 (Red)	M	2844.2	2841.7 (Disulfide bond formation between the cysteines)
Figure 2.14 (Blank)	M - 5 H_2O	2754.1	2754.3
	M - 4 H_2O	2772.2	2772.3
	M - 3 H_2O	2790.2	2790.1
	M - 4 H_2O + 1 IAA	2829.2	2829.4
	M - 3 H_2O + 1 IAA	2847.2	2847.3
Figure 2.14 (Red)	M + 3 IAA	3015.3	3016.0

2.5 REFERENCES

1. Li, B., Sher, D., Kelly, L., Shi, Y., Huang, K., Knerr, P. J., Joewono, I., Rusch, D., Chisholm, S. W., and van der Donk, W. A. (2010) Catalytic promiscuity in the biosynthesis of cyclic peptide secondary metabolites in planktonic marine cyanobacteria, *Proc. Natl. Acad. Sci. U.S.A. 107*, 10430-10435.

2. Zhang, Q., Yang, X., Wang, H., and van der Donk, W. A. (2014) High divergence of the precursor peptides in combinatorial lanthipeptide biosynthesis, *ACS Chem. Biol. 9*, 2686-2694.

3. Tang, W., and van der Donk, W. A. (2012) Structural characterization of four prochlorosins: a novel class of lantipeptides produced by planktonic marine cyanobacteria, *Biochemistry 51*, 4271-4279.

4. Yu, Y., Zhang, Q., and van der Donk, W. A. (2013) Insights into the evolution of lanthipeptide biosynthesis, *Protein Sci. 22*, 1478-1489.

5. Oman, T. J., Knerr, P. J., Bindman, N. A., Velasquez, J. E., and van der Donk, W. A. (2012) An engineered lantibiotic synthetase that does not require a leader peptide on its substrate, *J. Am. Chem. Soc. 134*, 6952–6955.

6. Shi, Y., Yang, X., Garg, N., and van der Donk, W. A. (2011) Production of lantipeptides in *Escherichia coli*, *J. Am. Chem. Soc. 133*, 2338-2341.

7. Lubelski, J., Rink, R., Khusainov, R., Moll, G. N., and Kuipers, O. P. (2008) Biosynthesis, immunity, regulation, mode of action and engineering of the model lantibiotic nisin, *Cell. Mol. Life Sci. 65*, 455-476.

8. Ekkelenkamp, M. B., Hanssen, M., Danny Hsu, S. T., de Jong, A., Milatovic, D., Verhoef, J., and van Nuland, N. A. (2005) Isolation and structural characterization of epilancin 15X, a novel lantibiotic from a clinical strain of Staphylococcus epidermidis, *FEBS Lett. 579*, 1917-1922.

9. Xie, L., Miller, L. M., Chatterjee, C., Averin, O., Kelleher, N. L., and van der Donk, W. A. (2004) Lacticin 481: in vitro reconstitution of lantibiotic synthetase activity, *Science 303*, 679-681.

10. Barber, M., Elliot, G. J., Bordoli, R. S., Green, B. N., and Bycroft, B. W. (1988) Confirmation of the structure of nisin and its major degradation product by FAB-MS and FAB-MS/MS, *Experientia 44*, 266-270.

11. Cheigh, C. I., Park, H., Choi, H. J., and Pyun, Y. R. (2005) Enhanced nisin production by increasing genes involved in nisin Z biosynthesis in Lactococcus lactis subsp. lactis A164, *Biotechnol. Lett. 27*, 155-160.

12. Velásquez, J. E., Zhang, X., and van der Donk, W. A. (2011) Biosynthesis of the

Antimicrobial Peptide Epilancin 15X and its Unusual N-terminal Lactate Moiety, *Chem. Biol. 18*, 857-867.

13. Zhang, Q., Yu, Y., Velasquez, J. E., and van der Donk, W. A. (2012) Evolution of lanthipeptide synthetases, *Proc. Natl. Acad. Sci. U.S.A. 109*, 18361-18366.

14. Knerr, P. J., and van der Donk, W. A. (2012) Discovery, biosynthesis, and engineering of lantipeptides, *Annu. Rev. Biochem. 81*, 479-505.

15. Oman, T. J., and van der Donk, W. A. (2010) Follow the leader: the use of leader peptides to guide natural product biosynthesis, *Nat. Chem. Biol. 6*, 9-18.

16. Kluskens, L. D., Kuipers, A., Rink, R., de Boef, E., Fekken, S., Driessen, A. J., Kuipers, O. P., and Moll, G. N. (2005) Post-translational Modification of Therapeutic Peptides By NisB, the Dehydratase of the Lantibiotic Nisin, *Biochemistry 44*, 12827-12834.

17. Chatterjee, C., Patton, G. C., Cooper, L., Paul, M., and van der Donk, W. A. (2006) Engineering dehydro amino acids and thioethers into peptides using lacticin 481 synthetase, *Chem. Biol. 13*, 1109-1117.

18. Levengood, M. R., and van der Donk, W. A. (2008) Use of Lantibiotic Synthetases for the Preparation of Bioactive Constrained Peptides, *Bioorg. Med. Chem. Lett. 18*, 3025-3028.

19. Kuipers, O. P., Bierbaum, G., Ottenwälder, B., Dodd, H. M., Horn, N., Metzger, J., Kupke, T., Gnau, V., Bongers, R., van den Bogaard, P., Kosters, H., Rollema, H. S., de Vos, W. M., Siezen, R. J., Jung, G., Götz, F., Sahl, H. G., and Gasson, M. J. (1996) Protein engineering of lantibiotics, *Antonie van Leeuwenhoek 69*, 161-169.

20. Cortés, J., Appleyard, A. N., and Dawson, M. J. (2009) Whole-cell generation of lantibiotic variants, *Methods Enzymol. 458*, 559-574.

21. Chatterjee, C., Paul, M., Xie, L., and van der Donk, W. A. (2005) Biosynthesis and mode of action of lantibiotics, *Chem. Rev. 105*, 633-684.

22. Rink, R., Kluskens, L. D., Kuipers, A., Driessen, A. J., Kuipers, O. P., and Moll, G. N. (2007) NisC, the Cyclase of the Lantibiotic Nisin, Can Catalyze Cyclization of Designed Nonlantibiotic Peptides, *Biochemistry 46*, 13179-13189.

23. Cotter, P. D., Deegan, L. H., Lawton, E. M., Draper, L. A., O'Connor, P. M., Hill, C., and Ross, R. P. (2006) Complete alanine scanning of the two-component lantibiotic lacticin 3147: generating a blueprint for rational drug design, *Mol. Microbiol. 62*, 735-747.

24. Appleyard, A. N., Choi, S., Read, D. M., Lightfoot, A., Boakes, S., Hoffmann, A., Chopra, I., Bierbaum, G., Rudd, B. A., Dawson, M. J., and Cortés, J. (2009) Dissecting structural and functional diversity of the lantibiotic mersacidin, *Chem. Biol. 16*, 490-

498.

25. Liu, W., and Hansen, J. N. (1992) Enhancement of the chemical and antimicrobial properties of subtilin by site-directed mutagenesis, *J. Biol. Chem. 267*, 25078-25085.

26. Field, D., Connor, P. M., Cotter, P. D., Hill, C., and Ross, R. P. (2008) The generation of nisin variants with enhanced activity against specific gram-positive pathogens, *Mol. Microbiol. 69*, 218-230.

27. Levengood, M. R., Knerr, P. J., Oman, T. J., and van der Donk, W. A. (2009) In vitro mutasynthesis of lantibiotic analogues containing nonproteinogenic amino acids, *J. Am. Chem. Soc. 131*, 12024-12025.

28. Xie, L., and van der Donk, W. A. (2004) Post-Translational Modifications during Lantibiotic Biosynthesis, *Curr. Opin. Chem. Biol. 8*, 498-507.

29. Dufour, A., Hindré, T., Haras, D., and Le Pennec, J. P. (2007) The biology of the lantibiotics of the lacticin 481 subgroup is coming of age, *FEMS Microbiol. Rev. 31*, 134-167.

30. Haft, D. H., Basu, M. K., and Mitchell, D. A. (2010) Expansion of ribosomally produced natural products: a nitrile hydratase- and Nif11-related precursor family, *BMC Biol. 8*, 70.

31. Hsu, S. T., Breukink, E., Tischenko, E., Lutters, M. A., De Kruijff, B., Kaptein, R., Bonvin, A. M., and Van Nuland, N. A. (2004) The nisin-lipid II complex reveals a pyrophosphate cage that provides a blueprint for novel antibiotics, *Nat. Struct. Mol. Biol. 11*, 963-967.

32. Hosoda, K., Ohya, M., Kohno, T., Maeda, T., Endo, S., and Wakamatsu, K. (1996) Structure determination of an immunopotentiator peptide, cinnamycin, complexed with lysophosphatidylethanolamine by ^{1}H-NMR, *J. Biochem. 119*, 226-230.

33. Tang, W., and van der Donk, W. A. (2013) The sequence of the enterococcal cytolysin imparts unusual lanthionine stereochemistry, *Nat. Chem. Biol. 9*, 157-159.

CHAPTER III: MECHANISTIC STUDIES OF CLASS II LANTHIPEPTIDE SYNTHETASES*

3.1 INTRODUCTION

As introduced in chapter II, ProcM is a member of the class II bifunctional lanthipeptide synthetases (generically termed LanMs).[1] ProcM is the only enzyme responsible for the dehydration and cyclization of 30 different substrates.[1, 2] Similar class II systems with a single enzyme and multiple, diverse substrate peptides are also found in other cyanobacteria.[2] In contrast, most LanM enzymes that have been investigated have only a single substrate.[3] The molecular details that allow this high substrate tolerance are largely unknown. Previous mechanistic studies on ProcM[4, 5] and NMR analysis of purified prochlorosin products[6] support enzymatic formation of all the thioether rings and showed that specific structures were generated for several prochlorosins. However, whether ProcM forms only one product for each substrate, or whether multiple products were formed and only one isomer was purified and structurally characterized was unclear prior to the work presented in this chapter. Indeed, cyclization of substrates that contain multiple cysteines and dehydro amino acids could result in products with alternative ring topologies (Figure 3.1).

---

*Reproduced in part with permission from Yu, Y., Mukherjee, S., and van der Donk, W. A. (2015) Product Formation by the Promiscuous Lanthipeptide Synthetase ProcM is under Kinetic Control, *J. Am. Chem. Soc. 137*, 5140-5148.

Using liquid chromatography electrospray ionization mass spectrometry (ESI-LC-MS), I show here that indeed ProcM is highly selective for the formation of one product from each substrate peptide, despite their highly diverse sequences. These remarkable observations prompted further investigations of the factors that contribute to both substrate tolerance and product selectivity.

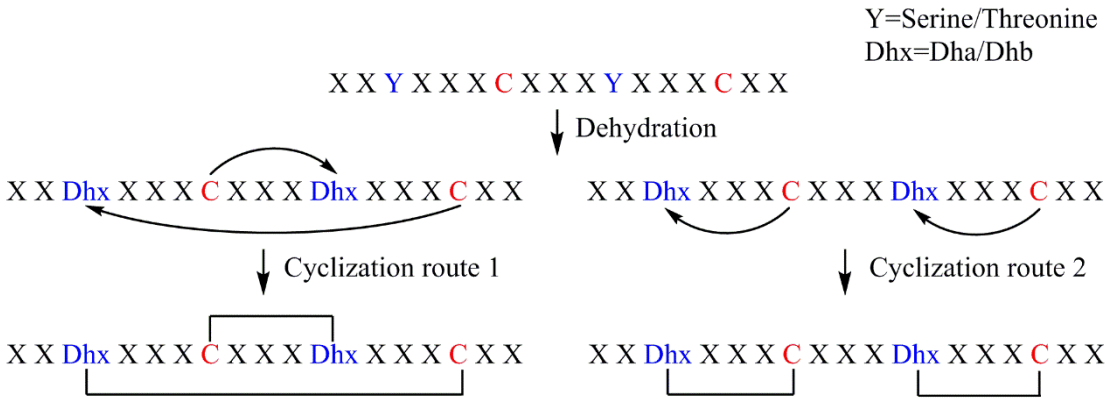

Figure 3.1. A peptide that contains two or more Cys residues and two or more dehydroamino acids can form multiple isomeric products as illustrated for a hypothetical example of a peptide with two cysteines and two dehydroamino acids.

The C-terminal domains of LanM enzymes have low but detectable sequence identity with class I cyclases (generically termed LanCs)[7] and the residues that are important for cyclization are strictly conserved based on structural[8] and mutagenesis studies[7] with the prototype NisC. For NisC, these residues include one histidine and two cysteines that serve as ligands to an active site zinc ion. These residues are also conserved in the cyclization domain of LanM

enzymes, and their importance has been experimentally confirmed in the case of the lacticin 481 synthetase LctM.[9] Surprisingly, a sequence alignment of ProcM with other LanM and LanC proteins (Figure 3.2) demonstrates that ProcM has three cysteines as the predicted zinc ligands. ProcM and its close homologues with three zinc-binding cysteines form a distinct clade in the phylogenetic tree of class II LanM synthetases,[10] and generally (but not always) are associated with multiple substrate peptides with diverse sequences.[2] As demonstrated in previous studies of other enzymes and model complexes, adding an additional sulfur-based ligand increases the reactivity of thiolates bound to a zinc ion.[11-15] Thus, the three-cysteine ligand set may be responsible for the promiscuous catalytic ability of ProcM and its homologues. Herein, the importance of these active site residues was investigated by mutagenesis, and the cyclization activity and regioselectivity of ring formation were determined. The role of the unique cysteine in the ProcM clade (Cys971 in ProcM) was assessed for the first time in this work, which provides important information to understand the unusual substrate tolerance and selectivity.

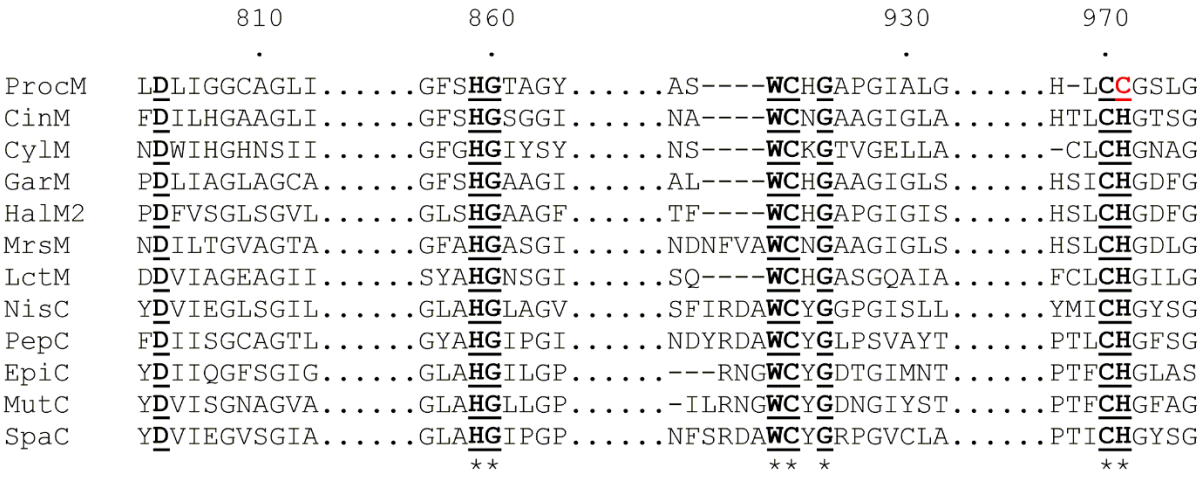

Figure 3.2. Sequence alignment of cyclase domains of selected LanMs with selected LanCs. Active site residues that are conserved among the LanC and LanM proteins are in boldface underlined type. The unique zinc-binding cysteine in ProcM is shown in red.

To understand better how ProcM can display both catalytic promiscuity and high specificity, several possible explanations were tested in this study. One possibility involves a highly organized order of the dehydration and cyclization reactions. To examine this hypothesis, the cyclization process was studied in isolation using dehydrated substrate and only the ProcM cyclase domain. A second recently suggested hypothesis to explain the properties of ProcM involves thermodynamic control of ring formation resulting in the accumulation of the most stable product. This hypothesis was born out of the observation that both NisC and the class II haloduracin synthetase HalM2 were able to efficiently catalyze retro-Michael additions on their products.[16] Thermodynamic control not only requires that the enzymes reversibly catalyze the Michael additions found in their products, but also can correct alternate cyclization patterns. This hypothesis was tested by incubation of ProcM with peptides containing incorrect rings

and by searching for incorrect intermediates formed during catalysis by ProcM. A third hypothesis involves kinetic control of ring formation, with the earlier cyclization events determining the regioselectivity of later cyclizations. Our investigations on ProcM reported here support this last model providing important new insights into the enzymatic mechanism of this unique clade of substrate tolerant class II lanthipeptide synthetases.

3.2 RESULTS

3.2.1 The paradoxical catalytic promiscuity and high selectivity of ProcM

For most of the lanthipeptides discovered to date, the thioether rings are installed by nucleophilic attack of Cys onto Dha or Dhb residues located N-terminal to the Cys.[17] But for prochlorosins, the thioether rings are formed by the addition of Cys residues to Dha or Dhb that are localized either N- or C-terminal to the Cys.[1] Five ProcA precursor peptides (ProcA1.1-G−1E, ProcA2.8, ProcA2.11, ProcA3.2, and ProcA3.3) were chosen for investigation out of the 30 possible substrates to provide a representative set of different ring topologies, ring size, and direction of cyclization (Figure 3.3). For Pcn1.1 (prochlorosin formed from ProcA1.1) and Pcn2.8, structures with two non-overlapping rings are formed, and for Pcn2.11, Pcn3.2, and Pcn3.3, structures with overlapping rings are generated.[1] The precursor peptides of these five prochlorosins containing an N-terminal hexa-histidine tag were co-expressed with ProcM in *E. coli,* and the modified peptides were purified by Ni^{2+} affinity chromatography and digested with endoproteases to remove all or most residues of the leader peptides. Samples were treated with iodoacetamide (IAA) to test for any partially cyclized products by reaction with free thiols

(Figure 3.4), and the resulting peptides were analyzed by ESI-LC-MS experiments. For all five peptides, the extracted ion chromatograms of fully dehydrated and cyclized products showed formation of one major product (Figure 3.3); any products with alternative ring patterns were only present in less than 5% of the major product. Tandem ESI-MS confirmed that the structures of the products were consistent with those determined in previous reports (Figure 3.5). The presence of y”9, y”8, y’’5 ions in the tandem MS spectrum for ProcA1.1 G-1E, and y”9, b39, y”8 ions for ProcA2.8 indicates these peptides contain two non-overlapping rings, whereas the absence of any y”12 through y”3 ions for ProcA3.3 strongly suggests that overlapping structures are present. These findings are also in full agreement with the structures determined by NMR spectroscopy.[6] Hence, this analysis confirmed that ProcM is both remarkably promiscuous with respect to its substrates and at the same time curiously specific in its product formation.

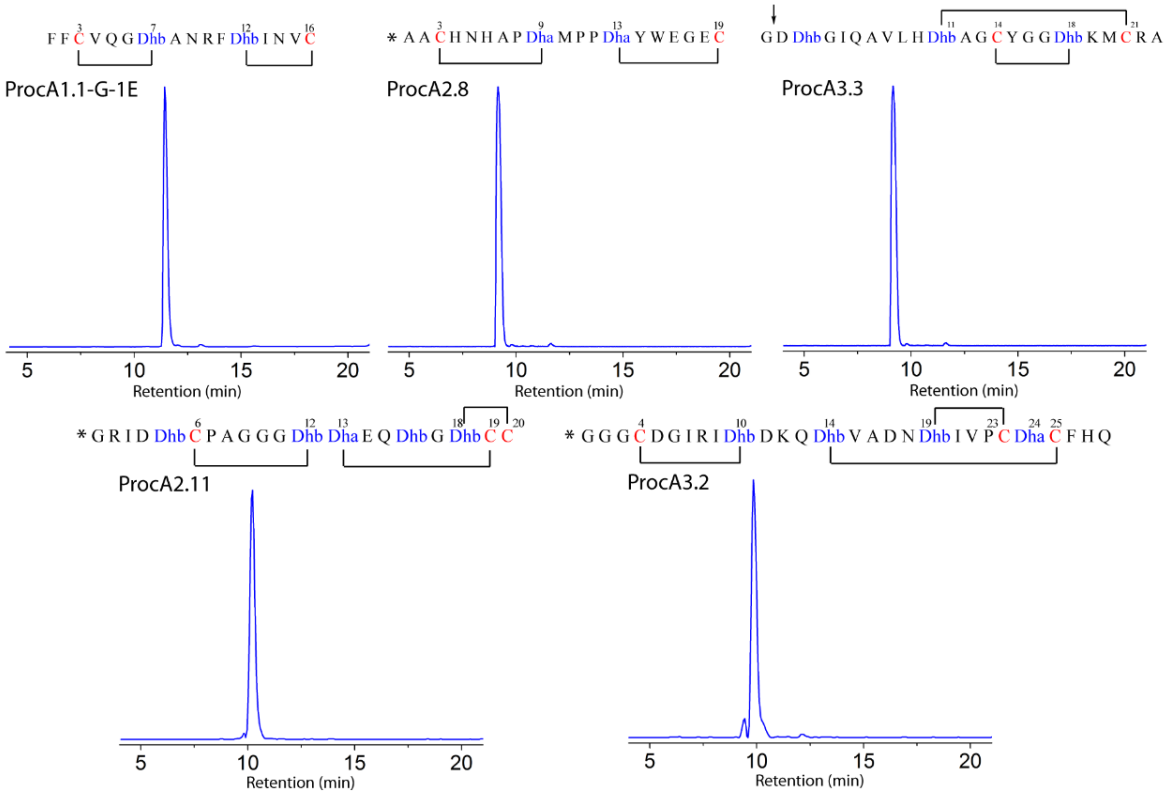

Figure 3.3. Extracted ion chromatograms (EICs) of the ESI-LC-MS experiments showing the fully dehydrated core-containing peptide ions from ProcA peptides co-expressed with ProcM, followed by digestion by endoproteases and treatment with IAA: ProcA1.1-G−1E and ProcA3.2 digested by GluC, ProcA2.11 digested by LysC, ProcA2.8 digested by trypsin and ProcA3.3 digested by AspN. The star symbol represents a short peptide sequence remaining from the leader peptide after treatment with the commercial proteases: ProcA2.8 (*=DELEGVAGG), ProcA2.11 (*=EDLNSHRQTLSEDELESVAGG), and ProcA3.2 (*=GVAGG). Arrow symbol stands for the site of AspN cleavage in ProcA3.3 core peptide.

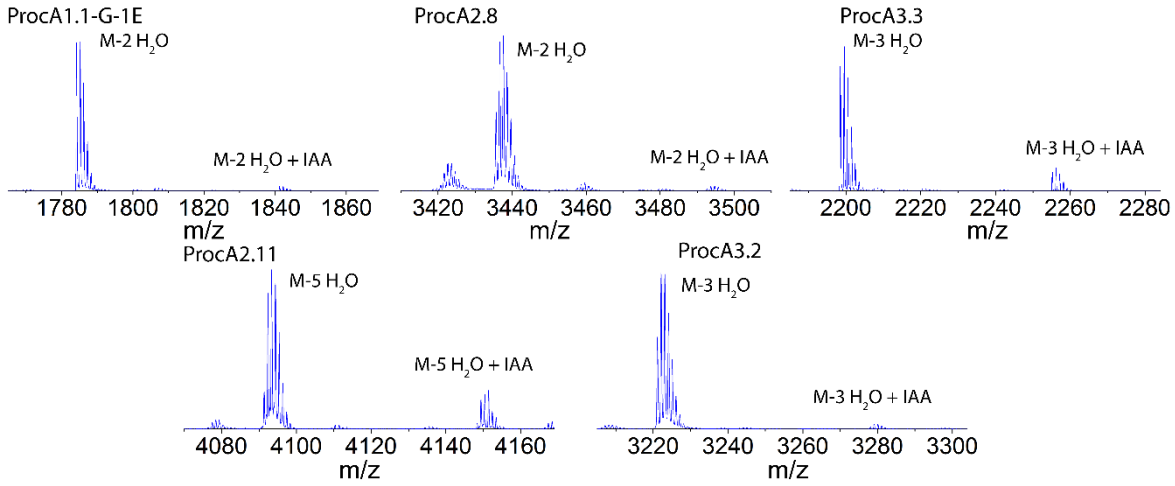

Figure 3.4. MALDI-TOF mass spectra of ProcA peptides co-expressed with ProcM, followed by digestion by endoproteases and treatment with IAA: ProcA1.1-G−1E and ProcA3.2 by GluC, ProcA2.11 digested by LysC, ProcA2.8 digested by trypsin and ProcA3.3 digested by AspN. Calculated and observed masses are shown in Table 3.3.

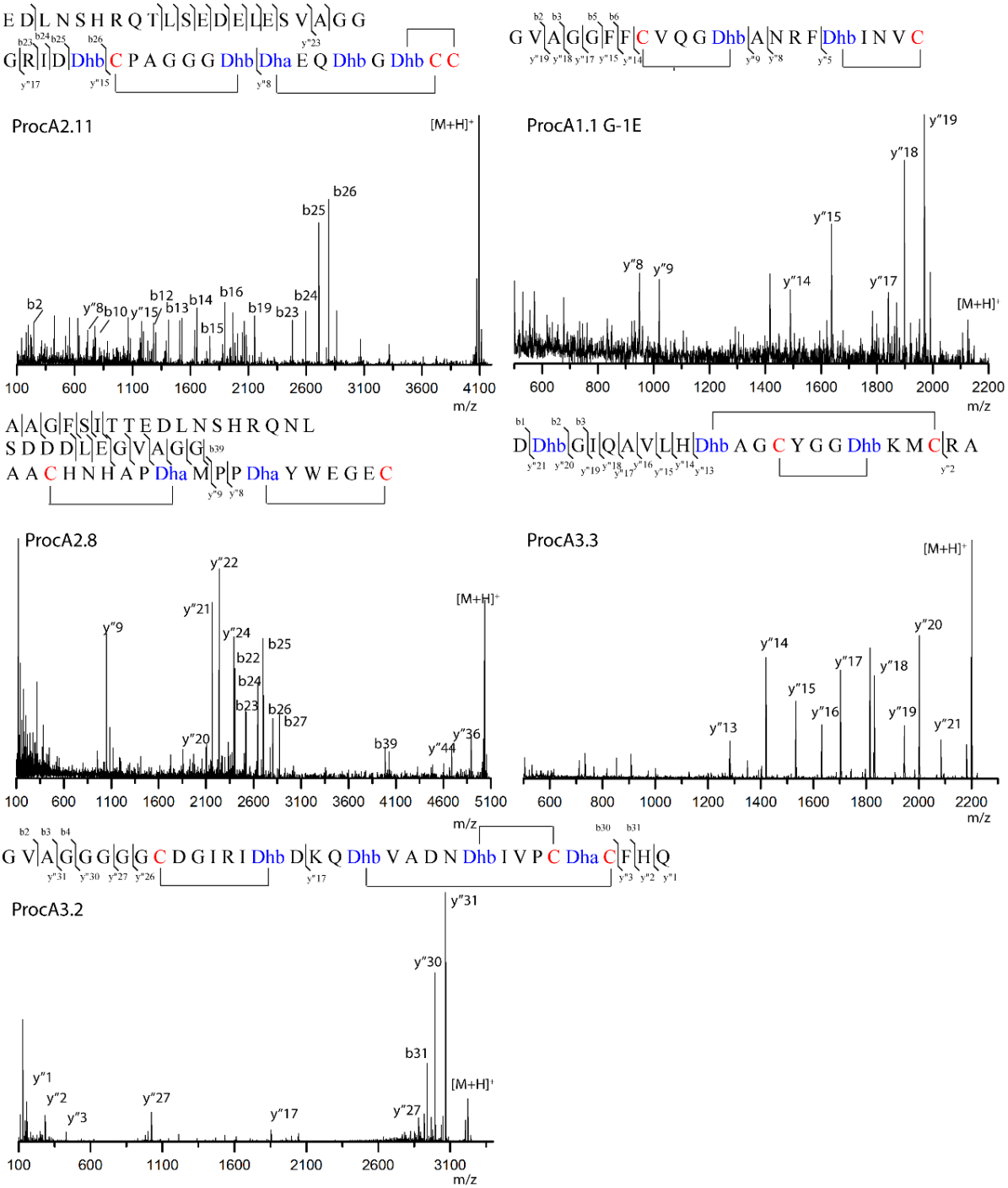

Figure 3.5. Tandem ESI-MS of the fully dehydrated core-containing peptide ions from ProcA peptides co-expressed with ProcM, followed by digestion by endoproteases and treatment with IAA: ProcA1.1-G−E and ProcA3.2 digested by GluC, ProcA2.8 and ProcA2.11 digested by LysC and ProcA3.3 digested by AspN.

3.2.2 The three zinc-binding cysteines are important for cyclization

Unlike other lanthipeptide cyclases that utilize two Cys and one His as zinc ligands, ProcM contains a Cys instead of a His at position 971, suggesting that it likely uses three Cys residues (Cys924, Cys970, and Cys971) for binding of the active site zinc ion. To confirm that these three Cys residues are important for zinc binding and for the cyclization activity of ProcM, a series of individual mutants were generated (C924A, C970A, C971A and C971H). The zinc content of WT ProcM and mutant proteins was measured by a spectrophotometric assay using the metallochromic indicator 4-(2-pyridylazo)resorcinol (PAR).[18] The proteins were expressed in *E. coli*, and then analyzed after purification and exhaustive dialysis in chelex-treated buffer. The resulting samples were treated with guanidine hydrochloride and 5,5´-dithio-bis(2-nitrobenzoic acid) (DTNB) to oxidize reactive cysteines to disulfides as previously described.[19] The orange $Zn(PAR)_2$ complex was formed after the addition of PAR and was quantified by the absorbance at 500 nm. Wild type (WT) ProcM and ProcM-C971H contained nearly stoichiometric quantities of zinc, while the relative zinc content for ProcM-C924A and ProcM-C970A was decreased by 50% (Figure 3.6). ProcM-C971A precipitated during the dialysis treatment in two independent attempts, thus zinc content could not be determined for this mutant.

Enzyme	Relative Zinc content
ProcM WT	1.00[a]
ProcM C924A	0.49
ProcM C970A	0.49
ProcM C971A	N/A
ProcM C971H	1.25

[a]Wild type ProcM contained 0.88 eq of zinc/protein.

Figure 3.6. Zinc content of WT ProcM and mutants.

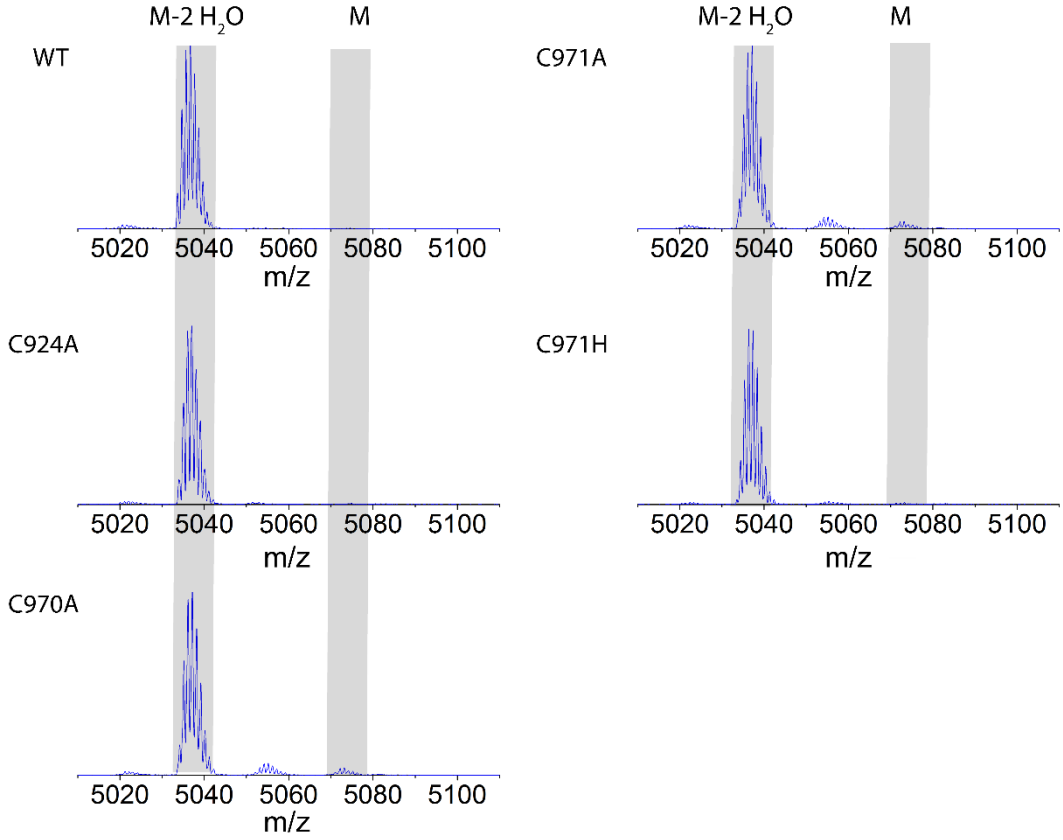

Figure 3.7. MALDI-TOF mass spectra of 25 μM ProcA2.8 peptide modified by 2 μM WT ProcM, ProcM-C924A, ProcM-C970A, ProcM-C971A, and ProcM-C971H *in vitro* for 1 h, followed by digestion by LysC.

The dehydration activities of WT ProcM and mutants were studied by *in vitro* enzymatic assays on the substrate ProcA2.8 followed by endoprotease LysC removal of most of the leader peptide and analysis by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS). As anticipated since the mutations are in the cyclization domain, the dehydration activities were not affected by these zinc ligand mutations (Figure 3.7). To study the cyclization activities, IAA was added to the LysC cleavage reaction to detect free cysteines in the peptides. With ProcA2.8 as substrate, all four Cys mutants showed a significant decrease in the cyclization activity compared to the WT enzyme, with the majority of the products carrying one IAA adduct, which indicated one ring was not formed (Figure 3.8). Considering about half of these mutants are apo-protein based on the zinc content analysis, these reactions were also performed with supplemented zinc or with four times the amount of enzyme, but similar results were observed. The partially cyclized peptides were analyzed by ESI-LC-MS of LysC-digested samples, which showed that it was the Cys3-Dha9 ring in ProcA2.8 that was formed slower for all four mutants (Figures 3.9 and 3.10, Table 3.4 lists the calculated and observed masses of each ion in the tandem MS analyses). Collectively, these experiments demonstrate that having three Cys ligands is important for the cyclization activity of ProcM. It is worthy to mention that assays using ProcM triple mutant (ProcM-C924A/C970A/C971A) resulted similar cyclization in ProcAs as the single mutants.

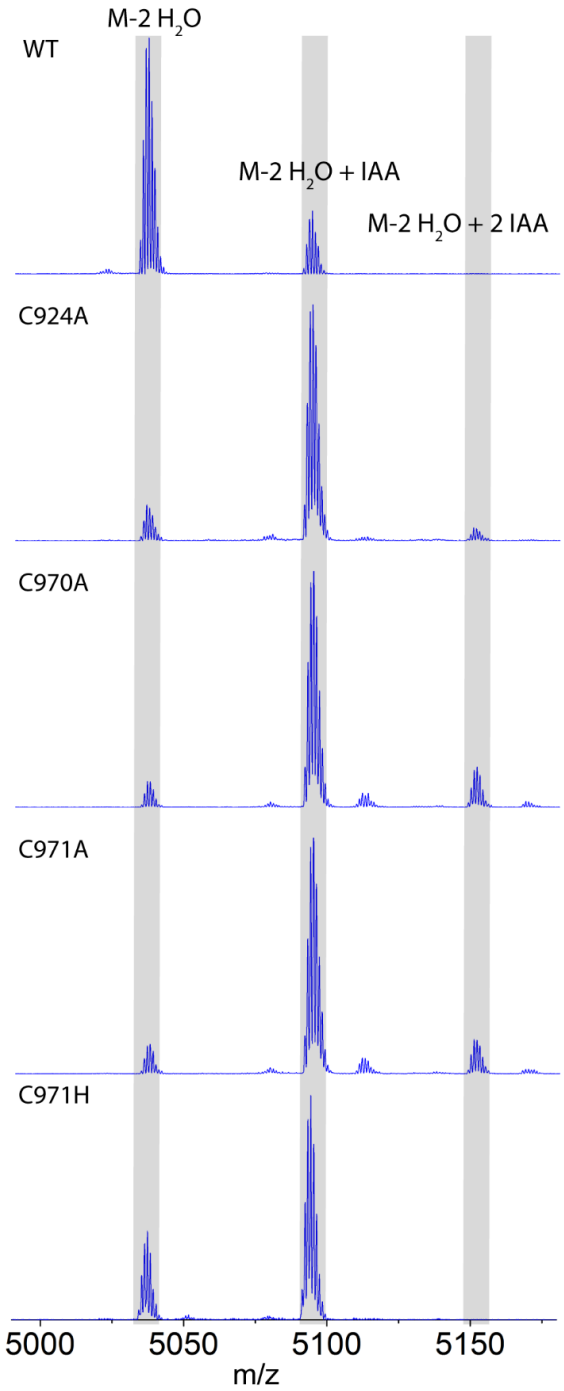

Figure 3.8. MALDI -TOF mass spectra of 25 μM ProcA2.8 peptide modified by 2 μM WT ProcM, ProcM-C924A, ProcM-C970A, ProcM-C971A, and ProcM-C971H *in vitro* for 1 h, followed by digestion by LysC and treatment with IAA.

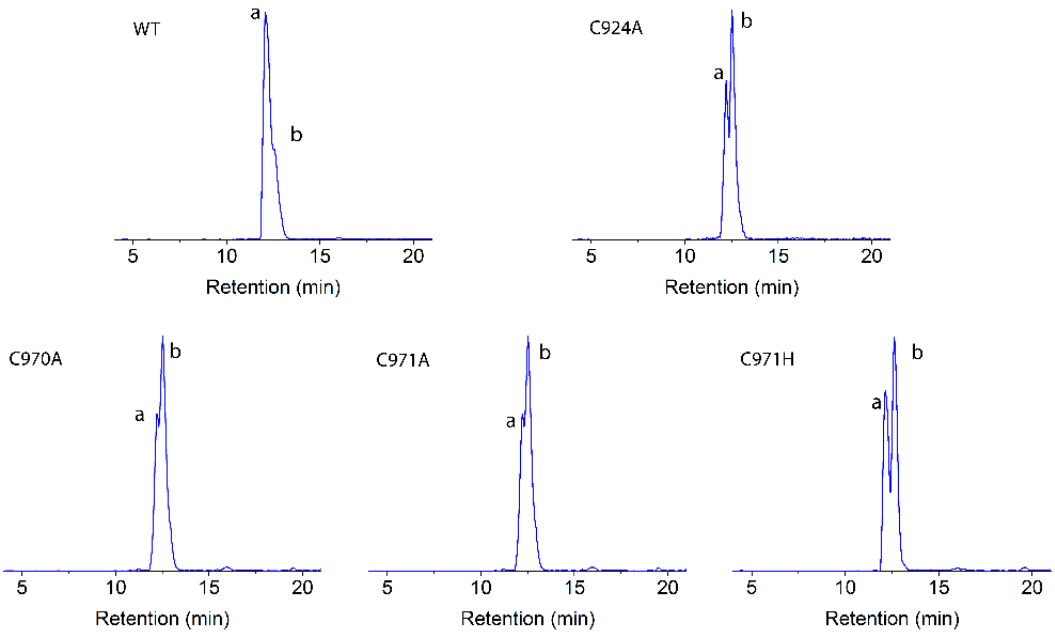

Figure 3.9. Extracted ion chromatograms (EICs) of the ESI-LC-MS experiments showing the fully dehydrated core-containing peptide ions of ProcA2.8 from Figure 3.7. Peptide **a** is fully cyclized, peptide **b** does not contain a ring between Cys3 and Dha9 (see Figure 3.10).

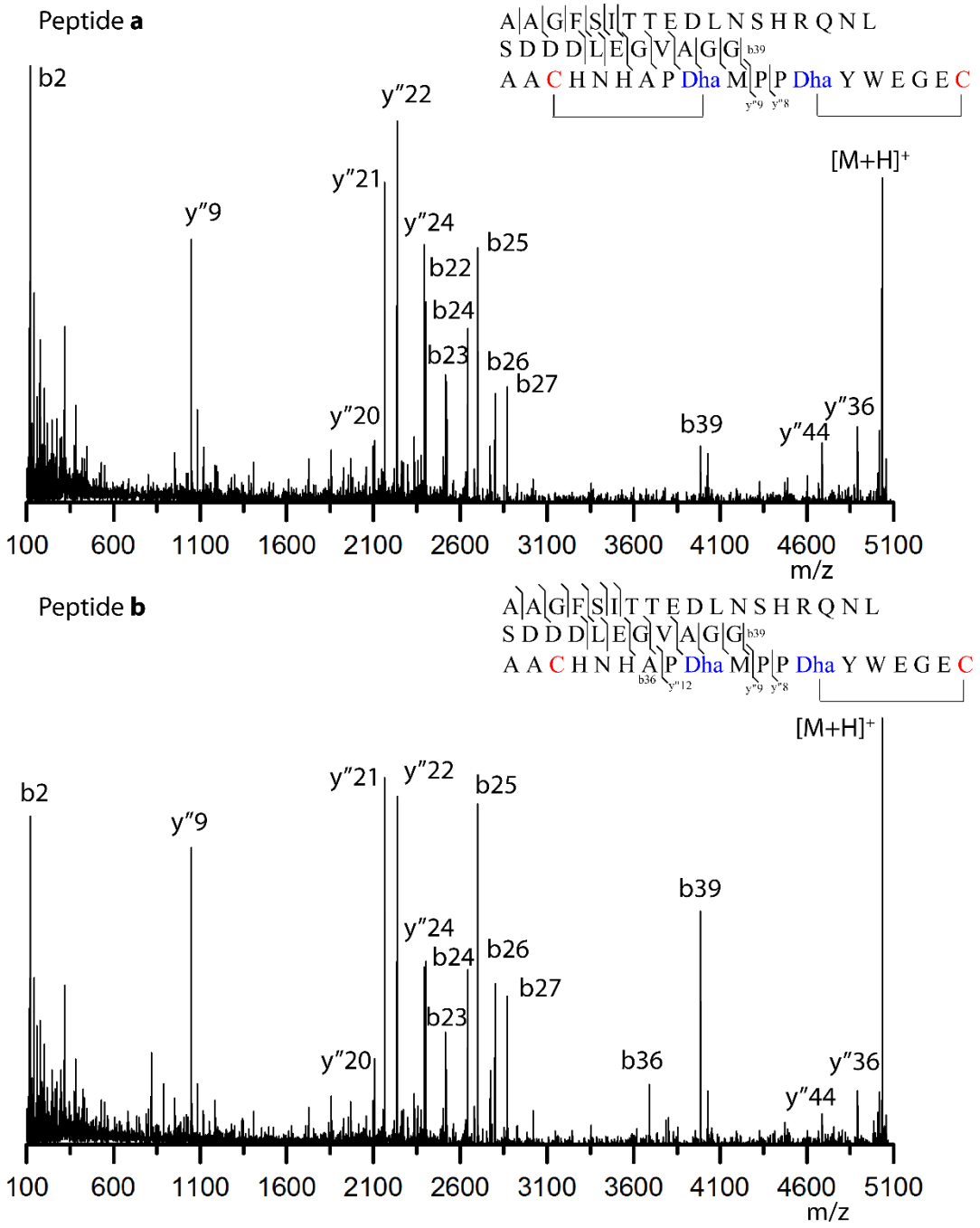

Figure 3.10. Tandem ESI-MS analysis of peptide **a** and peptide **b** shown in Figure 3.9. Peptide **a** is fully cyclized, peptide **b** does not contain a ring between Cys3 and Dha9. For calculated and observed masses, see Table 3.4.

3.2.3 Product distribution and cyclization rate of WT ProcM and ProcM mutants

Incubation of ProcA1.1 or ProcA2.8 with WT ProcM or ProcM-C971H *in vitro* resulted in formation of the same products as observed with WT ProcM (Figures 3.9 and 3.12). The rate of cyclization with these two substrates was studied *in vitro* by quenching the reaction at different time points. With WT ProcM, the first cyclization of Cys16 to Dhb12 in ProcA1.1 was almost completed in 30 min under the assay conditions used, and the second cyclization of Cys3 to Dhb7 was almost completed in 3 h (Figures 3.11 and 3.12, Table 3.5 lists the calculated and observed masses of each ion in the tandem MS analyses). With ProcM-C971H, the first cyclization of Cys16 to Dhb12 was nearly completed in 1 h, but the second cyclization was much slower, with less than 10% fully cyclized product observed at 3 h (Figure 3.11). The ratio was calculated using the ion counts for signals corresponding to the fully modified peptide and peptides with IAA adducts in the mass spectrum.

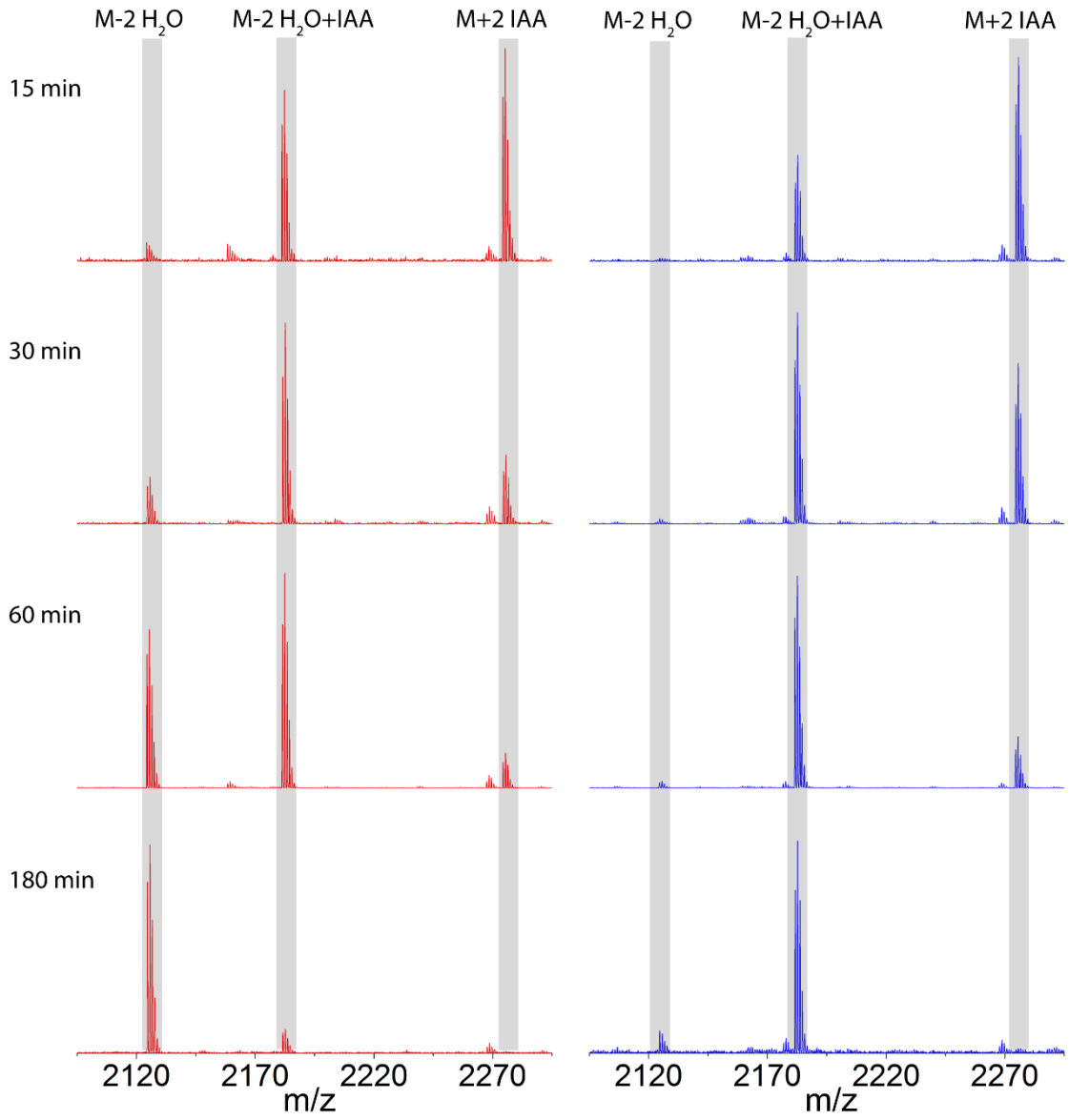

Figure 3.11. MALDI-TOF mass spectra of 10 μM ProcA1.1 modified by 0.2 μM WT ProcM (red) and ProcM-C971H (blue) for 15 min, 30 min, 60 min, and 180 min, followed by digestion by GluC and treatment with IAA.

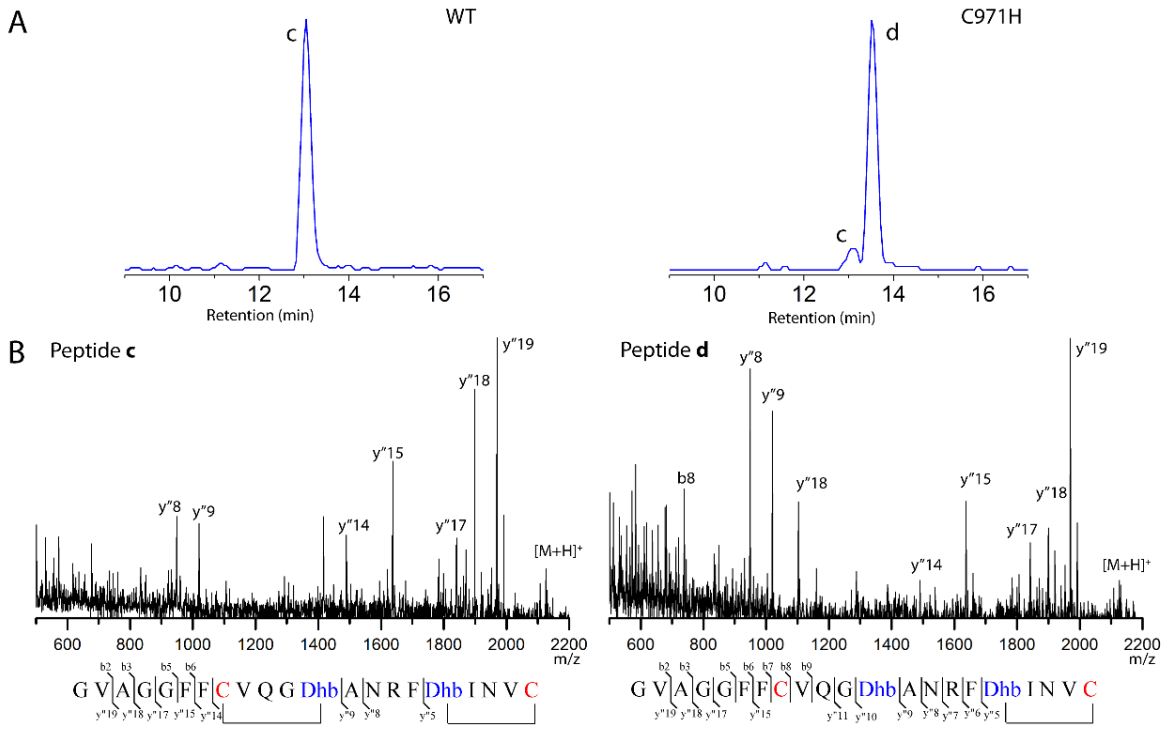

Figure 3.12. A. EICs of the ESI-LC-MS experiments showing the fully dehydrated core-containing peptide ions of ProcA1.1 from the reactions at 180 min from Figure 3. **B.** Tandem ESI-MS analysis of peptide **c** and peptide **d**. For calculated and observed masses, see Table 3.5.

The enzymatic assay with ProcA2.8 showed similar results (experiments performed by Subha Mukherjee). For WT ProcM, the first cyclization of Cys19 to Dha13 was completed in less than 15 min under the assay conditions used, and the second cyclization of Cys3 to Dha9 was almost completed in 1 h. With ProcM-C971H, the first cyclization of Cys19 to Dha13 was also completed in less than 15 min, but the rate of the second cyclization was decreased significantly, with <30% fully cyclized product observed after 90 min. By increasing the concentration of ProcM-C971H to 30 μM (with 50 μM peptide), the ratio of fully cyclized to partially cyclized ProcA2.8 after 1 h was increased (experiments performed by Subha

Mukherjee). This observation shows that ProcM-C971H was able to catalyze the second cyclization, although the rate was considerably slower compared to WT enzyme.

Surprisingly, for one of the investigated ProcAs, ProcA3.3, co-expression with ProcM-C971H in *E. coli* generated two different final products in contrast to WT ProcM, which made one specific product (peptide **1**, Figure 3.13A). The major product generated by ProcM-C971H (peptide **2**) was demonstrated by tandem ESI-MS to contain a novel non-overlapping ring topology (Figure 3.13B). ProcA3.3 was also incubated with ProcM-C971H and the other mutants *in vitro* for 3 h, and the product distribution was analyzed by ESI-LC-MS. The ring patterns of each product and intermediate were determined by tandem ESI-MS and verified by comparison to standards (Figure 3.14). Tandem mass spectra of each product and intermediate are shown in Figures 3.13 and 3.15, and Table 3.6 lists the calculated and observed masses of each ion in the tandem MS analyses. For all zinc-ligand mutants, peptide **2** was the major final product and peptide **5** was the major intermediate that leads to the formation of peptide **2** (Figures 3.14 and 3.16). Mutating the proposed active site residue H859 (Figure 3.2) to Ala resulted in a complete loss of the cyclization activity of ProcM (Figure 3.16), which is consistent with previous mutagenesis studies on NisC.[9] While mutating D804 to Asn, which is the residue that forms a hydrogen bond with the active site histidine residue in the NisC crystal structure,[11] resulted in a dramatic decrease in the cyclization rate but no changes in the regioselectivity of the product formation (Figure 3.16).

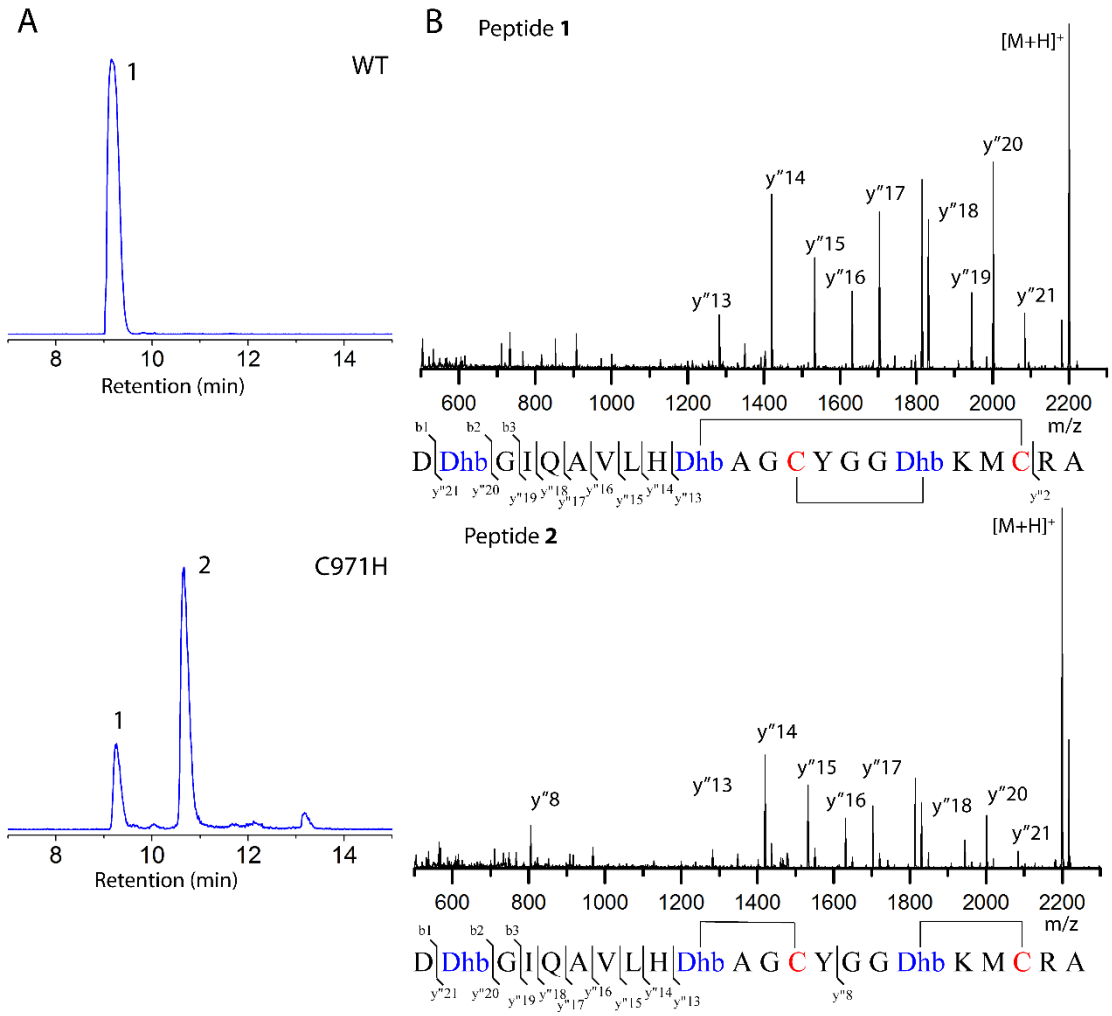

Figure 3.13. A. EICs of the ESI-LC-MS experiments showing the 3-fold dehydrated Δ1 core peptide ion (Δ1 because the first Gly is removed by AspN) produced by ProcA3.3 co-expressed with WT ProcM or ProcM-C971H, and subsequently digested with AspN and treated with IAA. **B.** Tandem ESI-MS of peptide **1** and peptide **2**. For calculated and observed masses, see Table 3.6.

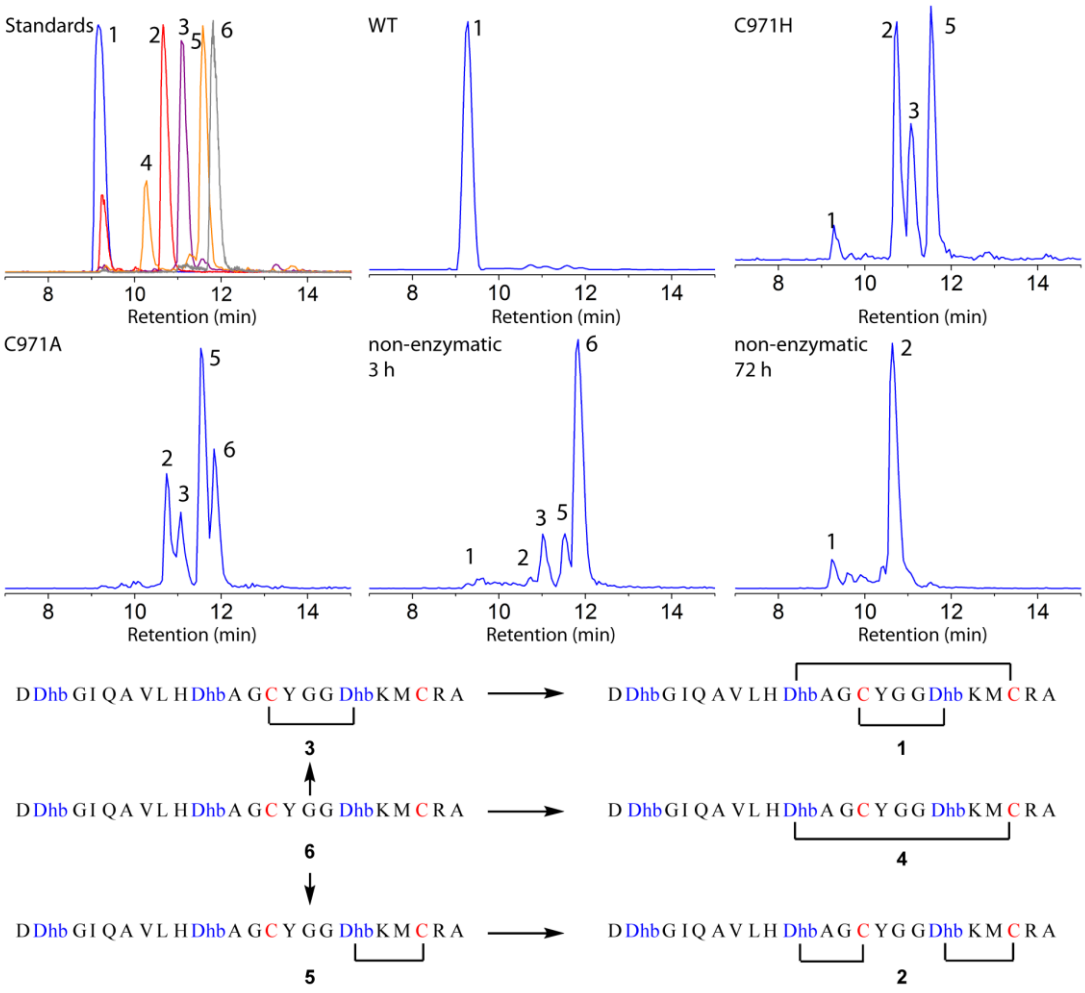

Figure 3.14. EICs of the ESI- LC-MS experiments showing the 3-fold dehydrated Δ1 core peptide ion obtained by AspN digestion of 25 μM ProcA3.3 incubated with 2 μM WT ProcM, ProcM-C971A, and ProcM C971H for 3 h. Standards: **1** (blue), ProcA3.3 co-expressed with WT ProcM; **1** and **2** (red), ProcA3.3 co-expressed with ProcM-C971H; **3** (purple), semi-synthetic ProcA3.3 containing a ring between Cys14 and Dhb18 (provided by Subha Mukherjee); **4** and **5** (orange), semi-synthetic ProcA3.3 with Cys14 protected that was modified by WT ProcM and subsequent deprotected (provided by Subha Mukherjee); and **6** (grey), dehydrated ProcA3.3. Non-enzymatic: dehydrated ProcA3.3 incubated with reaction

buffer at pH 8 for 3 h or 72 h. All the peptides were digested by AspN to remove the leader peptide.

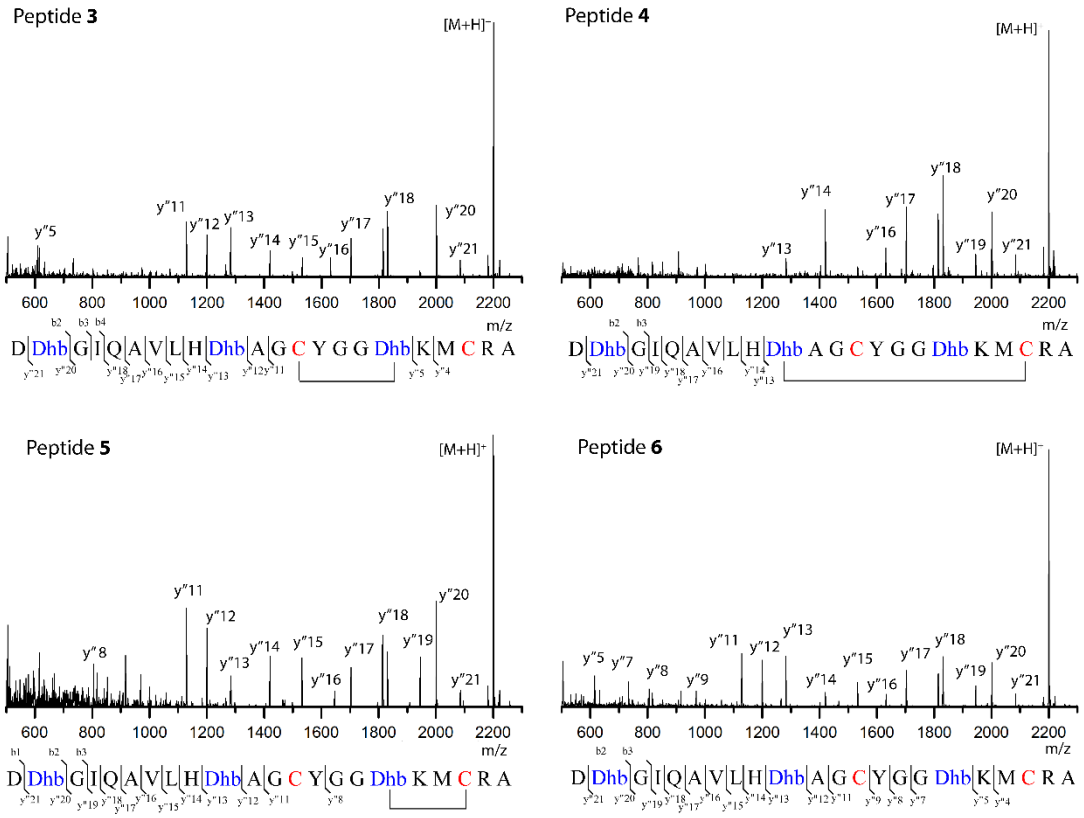

Figure 3.15. Tandem ESI-MS analysis of peptides **3**, **4**, **5** and **6** shown in Figure 3.14. For calculated and observed masses, see Table 3.6.

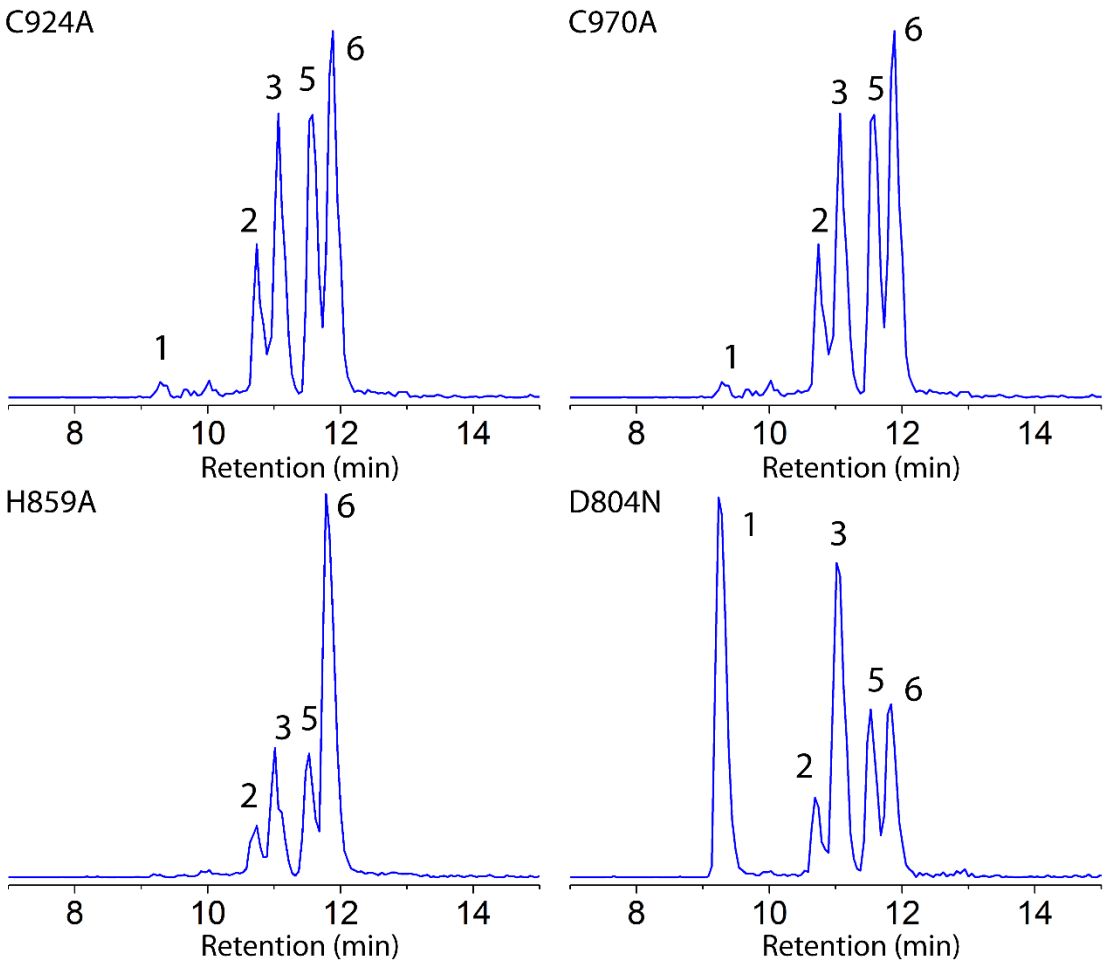

Figure 3.16. EICs of the ESI-LC-MS experiments showing the 3-fold dehydrated Δ1 core peptide ion from 25 μM ProcA3.3 incubated with 2 μM ProcM mutants for 3 h. All the reactions were treated by AspN prior to ESI-LC-MS analysis.

Given the interesting observation of a switch in product distribution with ProcA3.3, we also examined the time dependence of processing. Incubation of ProcA3.3 with WT ProcM resulted in near completion of the first cyclization of Cys14 to Dha18 in 15 min under the assay conditions, whereas the second cyclization of Cys21 to Dha11 took up to 90 min to complete (Experiments performed by Subha Mukherjee). For ProcM-C971H, the first cyclization of

Cys21 to Dha18 was nearly completed in 90 min, but the second cyclization of Cys14 to Dha11 was much slower, with less than 10% fully cyclized product observed at 90 min.

To determine the intrinsic regioselectivity of ring formation, non-enzymatic cyclization of ProcA3.3 was studied. First, the two cysteines in ProcA3.3 were linked by a disulfide bridge using oxidized glutathione, which resulted in the expected mass loss of two Da. The oxidized ProcA3.3 was dehydrated by ProcM *in vitro*, and the 3-fold dehydrated product was purified. Reduction by TCEP treatment shifted the mass of the peptide by two Da (Figure 3.17). Upon TCEP reduction, the non-enzymatic cyclization was investigated at the same pH of the enzymatic assay (pH 8) for 3 h. Very little cyclization occurred compared to the enzymatic reactions, illustrating that the mutants still accelerated the cyclization process (Figure 3.14). The regioselectivity of non-enzymatic cyclization was very similar to that of the zinc ligand mutants, with the predominant products being peptide **2** and intermediate **5** that leads to peptide **2** (Figure 3.14).

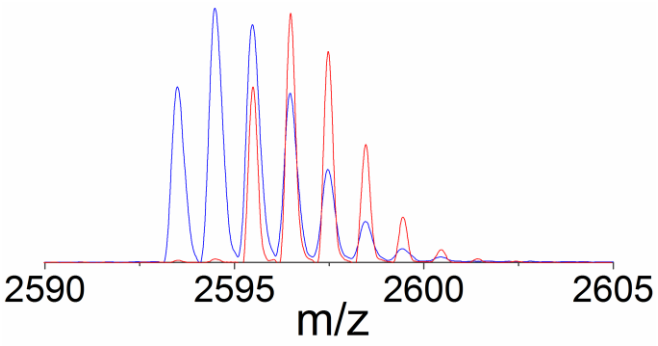

Figure 3.17. MALDI-TOF mass spectra of the GluC-cleaved dehydrated ProcA3.3 before (red) and after (blue) TCEP reduction.

3.2.4 The order of dehydration and cyclization does NOT determine the ring topology

It has been reported for nisin, haloduracin, and curvopeptin that the dehydration and cyclization reactions are alternating processes during biosynthesis.[5, 20-22] Thus, a possible explanation for the unusual selectivity of the ring formation in the structurally diverse prochlorosins is that the ring patterns are determined by the order of dehydration and cyclization. The order of dehydration has been established in a previous study for ProcA2.8 and ProcA3.3.[4] To study cyclization in isolation, the putative cyclase domains of WT ProcM and ProcM-C971H were obtained by expressing the C-terminal domains spanning residues 655-1068 (ProcM-655-1068 and ProcM-C971H-655-1068). The cyclization reaction catalyzed by ProcM-655-1068 was studied with dehydrated ProcA3.3 that was generated as described in the previous section. As will be shown in section 4.2.3, incubation of WT ProcM-655-1068 with fully dehydrated ProcA3.3 resulted in a similar product distribution as with the WT enzyme at a similar extent of conversion, strongly favoring peptide **1** and intermediate **3** leading to peptide **1**. Conversely, incubation of ProcM-C971H-655-1068 with fully dehydrated ProcA3.3 resulted in preference for intermediate **5** leading to peptide **2**. These results are not consistent with a tight regulation of the order of dehydration and cyclization as a determining factor for the selectivity for certain ring patterns in prochlorosins, since decoupling of the dehydration reaction from the cyclization events gave very similar product distributions compared to the selectivities observed with the full length proteins.

3.2.5 Thermodynamic control involving reversible ring formation is NOT observed

One attractive explanation for the catalytic selectivity of ProcM is thermodynamic control over the selectivity of ring formation. This model requires reversibility of the cyclization reaction to allow the generation of the most stable product. Indeed, reversible ring formation has recently been observed for the haloduracin synthetase HalM2 and the nisin cyclase NisC.[16] Thermodynamic control not only requires ring opening of the rings found in the final product, but also ring opening of energetically less favored "incorrectly" cyclized products and intermediates. At present, such activity has not been investigated for any lanthipeptide. To probe this model for ProcM, three different experiments were carried out. First, the non-physiological product **2** was obtained by co-expressing ProcA3.3 with ProcM-C971H and peptide **2** was subjected to the *in vitro* enzymatic assay with WT ProcM. Conversion of peptide **2** to peptide **1** was not observed (Figure 3.18A) Furthermore, ProcM-C971H did not convert peptide **1** to peptide **2** either (Figure 3.18B). However, it is possible that the final cyclized peptides were not accessible to the active site of the enzyme. Indeed, for all substrates investigated in this study, the second cyclization event was very slow, possibly reflecting decreasing affinity of the enzyme for the peptide as more rings are installed. Therefore, we also investigated potential reversibility on an incorrectly cyclized intermediate. A ProcA3.3 derivative with a Cys21-Dhb18 ring was generated by orthogonal protection of Cys14 (performed by Subha Mukherjee).[4] Incubation of the intermediate with WT ProcM did not result in conversion of this "wrong intermediate" to the correct product with overlapping rings,

but instead a product with non-overlapping rings was formed. In other words, once the incorrect first ring was installed, ProcM proceeded to the incorrect product.

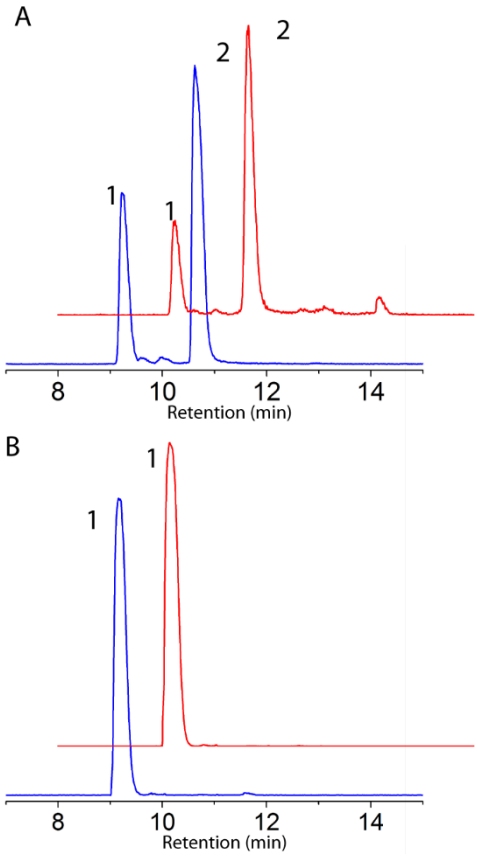

Figure 3.18. A. EICs of the ESI- LC-MS experiments showing the Δ1 core peptide ion from ProcA3.3 co-expressed with ProcM-C971H (red), and subsequently treated with WT ProcM (blue). **B.** EICs of the ESI-LC-MS experiments showing the Δ1 core peptide ion from ProcA3.3 co-expressed with WT ProcM (red), and subsequently treated with ProcM-C971H (blue). All peptides were digested by AspN and treated with IAA.

As a final test of potential thermodynamic control, the *in vitro* enzymatic assay of WT ProcM and ProcA3.3 was quenched at different time points and analyzed by ESI-LC-MS to identify the intermediates (Figure 3.19). At all time points, peptides **1** and **3** were detected at significant levels whereas incorrect intermediate peptide **5** was only observed at very low levels (possibly from background non-enzymatic cyclization). Collectively, these results argue against thermodynamic control of product formation for ProcM.

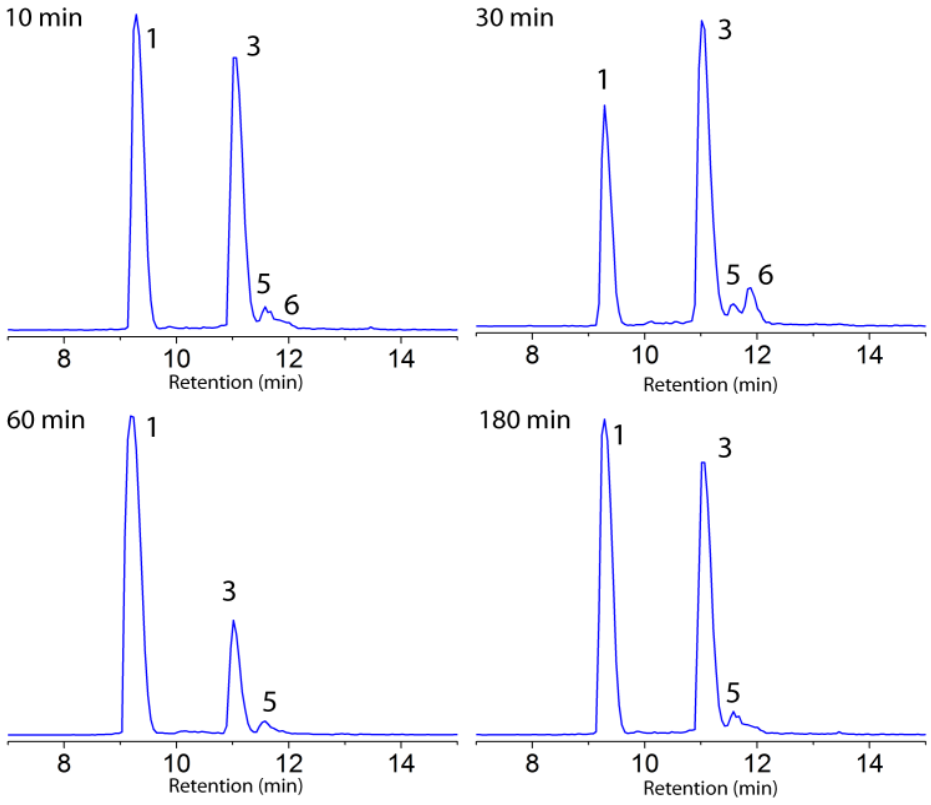

Figure 3.19. EICs of the ESI-LC-MS experiments showing the 3-fold dehydrated Δ1 core peptide ion from ProcA3.3 incubated with WT ProcM for 15 min, 30 min, 1 h, and 3 h. The reactions were digested by AspN.

3.3 DISCUSSION AND OUTLOOK

In this study, I demonstrated that, curiously, the substrate promiscuity of ProcM does not come with a cost of reduced product selectivity as for all the peptides investigated, the enzyme makes essentially single products, with only very minor amounts, if any, of alternative ring topologies observed. We investigated several possibilities for the determinants of the observed selectivity for certain ring topologies. In a previous study, we ruled out that the high selectivity for the final ring topology is a consequence of non-enzymatic processes.[4] Here we again show that non-enzymatic cyclization is much slower than the enzymatic process, and we show that the regioselectivity for non-enzymatic cyclization of ProcA3.3 is different than that observed with WT ProcM. Hence, the enzyme must exert a specific influence on the regiochemistry of cyclization and is not simply accelerating formation of a product that is already favored by the substrate. In this work we ruled out that the selectivity is caused by tight control between the dehydration and cyclization activities, or by thermodynamic control. A very recent study illustrated that two lantibiotic cyclization processes involved in the biosynthesis of the class I lantibiotic nisin and the class II lantibiotic haloduracin are reversible when the cyclization enzymes NisC and HalM2, respectively, were exposed to their final correctly cyclized products.[16] ProcM on the other hand appears much less reactive towards its products. All attempts to detect products of retro-Michael processes using the same methodology as used for NisC and HalM2,[16] whether with the native ProcM products or incorrectly cyclized products or intermediates, have been unsuccessful. No ring-opened products were detected with ProcM (~5% estimated detection limit), whereas NisC and HalM2 generated substantial amounts of

peptides in which up to five and four rings were opened, respectively.[16] These observations may be explained by the very different kinetic profiles of the enzymes.[5] The reactions that HalM2 catalyzes on its substrate HalA2 become progressively faster, whereas the opposite is observed here and in a previous study[5] for ProcM. A well-known problem in feeding intermediates to enzymes is that they cannot always access the catalytically competent form of their cognate enzyme, for instance if such an intermediate is normally processed further while still bound in the enzyme active site. However, ProcM is not processive in its catalysis with wild type substrates as shown here and in a previous study,[5] and hence this explanation for the lack of reversibility does not hold. Instead, we believe that the difference between HalM2/NisC and ProcM is the result of poor recognition of cyclized intermediates by the latter enzyme. We have suggested that the HalA2 peptide becomes increasingly complementary to the HalM2 active site as the posttranslational modifications take place, whereas ProcM, which makes 30 structurally very different products, cannot be complementary in shape and/or electrostatics to all its products.[5] Thus, the fully processed HalA2 peptide likely interacts much more favorably with HalM2 than the processed ProcA peptides interact with ProcM, explaining the observed reversibility for HalM2 and the apparent lack of reversibility for ProcM.

If neither the inherent reactivity, nor the thermodynamics of the products or tight coupling between dehydration and cyclization events determine the outcome of ProcM catalyzed cyclization reactions, what might govern the high observed regioselectivities? One possibility is that the enzyme-catalyzed selectivity for formation of the first ring determines the overall outcome. The mature products of ProcA1.1, 2.8, and 3.3 all have two rings in their final

products but with different ring sizes and topologies. With all three peptides, ProcM forms the first ring much faster than the second ring. Since we demonstrated that ProcM does not open up any formed rings, kinetic control over the regioselectivity of formation of the first ring determines the final ring topology for these peptides. The regioselectivity of the formation of that first ring is enzyme controlled because for ProcA3.3, non-enzymatic cyclization is both much slower and results in a different first ring compared to ProcM.

Zn^{2+} is often used in enzymes that catalyze alkylation of thiols, and it has been noted that these enzymes typically have a net negatively charged zinc site in their substrate-bound form (i.e. more than two negatively charged ligands from enzyme + substrate).[12] In addition, the negatively charged ligands in these enzymes typically are predominantly thiolates. It has been suggested that this ligand environment might either serve to facilitate substrate thiolate dissociation to release the nucleophile, or to increase the reactivity of the Zn-bound thiolate of the substrate in an associative process.[12] Although studies on model complexes have shed some light onto this question,[13, 14, 23, 24] at present, it is not clear whether the enzymes use a dissociative or associative mechanism. Therefore, the exact nature of substrate activation in lanthipeptide cyclases is not yet clear, but this study is in agreement with the notion that an increase in thiolate ligands increases the enzyme's activity.[15,16,19] Interestingly, all the zinc ligand mutants that perturb the three Cys ligand set (four Cys after binding of the substrate, Figure 3.20) display the same general regioselectivity as the non-enzymatic reaction. All these mutants also have attenuated cyclization activity. Hence, it appears that the unique three-cysteine ligand set not only increases the reactivity of the zinc site, but that this increased

reactivity also can overcome the inherent selectivity of the substrate. How this is achieved for a very diverse set of substrates is not clear, and likely involves substrate-enzyme interactions that await structural information on ProcM and its complexes with the substrates. What is clear is that the high reactivity afforded by the three-cysteine ligand set may be required for full cyclization of the prochlorosins at a reasonable rate since the second cyclization was very slow for all the mutant enzymes.

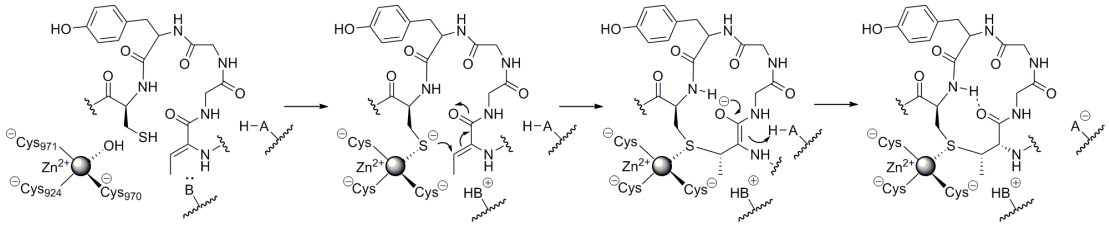

Figure 3.20. Proposed mechanism for the cyclization reaction catalyzed by ProcM illustrated for the formation of the Cys14 to Dhb18 ring in ProcA 3.3.

3.4 EXPERIMENTAL PROCEDURES

3.4.1 Construction of the expression plasmids for ProcM mutants

The generation of pRSF-Duet1-*procM* was described in section 2.4.3. The ProcM mutant constructs were generated by site-directed mutagenesis using pRSF-Duet1-*procM* (MCS1) as temperate and the primers in Table 3.1. The ProcM-C924A/C970A/C971A construct was generated from pRSF-Duet1-*procM*-C924A (MCS1) using *procM*_C970A/C971A_FP and *procM*_C970A/C971A_RP as primers. PCR amplifications were performed using Phusion polymerase by 30 cycles of denaturing (94 °C for 30 s), annealing (55 °C for 30 s), and

extending (72 °C for 3.5 min). The resulting PCR product was treated by DpnI, purified using Qiagen gel extraction kit, and used to transform *E. coli* DH5α. The sequences of the mutants were checked by DNA sequencing.

The genes encoding WT ProcM-655-1068 and ProcM-C971H-655-1068 were PCR amplified using Phusion polymerase by 30 cycles of denaturing (94 °C for 30 s), annealing (55 °C for 30 s), and extending (72 °C for 45 s) using *procM*_655_EcoRI_FP as forward primer, *procM*_Duet1_NotI_RP as reverse primer, and pRSF-Duet1-*procM* (MCS1) and pRSF-Duet1-*procM*-C971H (MCS1) as templates. The PCR products were digested with EcoRI and NotI restriction enzymes and ligated into the first multiple cloning site of a pRSF-Duet1. The sequence of the insert was checked by DNA sequencing.

Table 3.1. Primers used in section 3.4.1.

Primer Name	**Primer Sequences (5'-3')**
*procM*_C924A_FP	CTTCATGGCTAGCTGGGCCCATGGTGCCCCGGGTATTG
*procM*_C924A_RP	CAATACCCGGGGCACCATGGGCCCAGCTAGCCATGAAG
*procM*_C970A_FP	GTTTCGACGGATCATCTTGCCTGCGGATCACTGGGTTTGATG
*procM*_C970A_RP	CATCAAACCCAGTGATCCGCAGGCAAGATGATCCGTCGAAAC
*procM*_C971A_FP	GTTTCGACGGATCATCTTTGCGCCGGATCACTGGGTTTGATG
*procM*_C971A_RP	CATCAAACCCAGTGATCCGGCGCAAAGATGATCCGTCGAAAC
*procM*_C971H_FP	GTTTCGACGGATCATCTTTGCCACGGATCACTGGGTTTGATG
*procM*_C971H_RP	CATCAAACCCAGTGATCCGTGGCAAAGATGATCCGTCGAAAC
*procM*_Duet1_EcoRI_FP	CAGGATCCGAATTCGATGGAAAGTCCATCATCTTGG
*procM*_655_EcoRI_FP	GAATCCGAATTCGATGTCATCATCGGCTGAGACGA
*procM*_C970A/ C9701_FP	GTTTCGACGGATCATCTTGCCGCCGGATCACTGGGTTTGATG
*procM*_C970A/ C9701_RP	CATCAAACCCAGTGATCCGGCGGCAAGATGATCCGTCGAAAC

3.4.2 Construction of the co-expression plasmids for ProcAs and ProcM /ProcM mutants

The *procM* and *procM* mutant genes were amplified by PCR in 30 cycles of denaturing (98 °C for 30 s), annealing (55 °C for 30 s), and extending (72 °C for 2 min) with the plasmid pRSF-Duet1-*procM*/*procM* mutants as templates and using *procM*_Duet1_NdeI_FP as forward primer and *procM*_Duet1_KpnI_RP as reverse primer. The PCR products were digested with NdeI and KpnI restriction enzymes and ligated into the second multiple cloning site (MCS2) in the pRSF-Duet1 vector to generate the construct pRSF-Duet1-*procM*/*procM* mutant (MCS2). The sequences of the resulting products were confirmed by DNA sequencing.

ProcA genes were amplified by PCR using 30 cycles of denaturing (98 °C for 10 s), annealing (55 °C for 30 s), and extending (72 °C for 15 s) using the primers listed in Table 3.2 and the plasmids pET15b-*procA*1.1/*procA*2.8/*proc3.3* as templates.[1, 25] The resulting DNA inserts and the pRSF-Duet1 vector containing *procM*/*procM* mutants in MCS2 were digested with EcoRI and NotI and ligated. The sequences of the resulting products were confirmed by DNA sequencing.

Table 3.2. Primers used in section 3.4.2.

Primer Name	**Primer sequence (5'-3')**
*procM*_Duet1_NdeI_FP	GGTTGGTTCATATGGAAAGTCCATCATCTTGG
*procM*_Duet1_KpnI_RP	AAGTAGTTGGTACCTTATTCAGTAGGCCAGAGAC
*procA*1.1_EcoRI_FP	AAGAATTCGATGAAAAAGCGACTCAAC
*procA*1.1_NotI_RP	AAGCGGCCGCTCAGCACACATTGATAG
*procA*2.8_EcoRI_FP	GGTGAGTGGAATTCGATGTCAGAAGAGCAACTGAAGGC
*procA*2.8_NotI_RP	ATATAATGCGGCCGCTTAGCACTCACCCTC
*procA*3.3_EcoRI_FP	GGTGAGTGGAATTCGATGTCAGAAGAACAACTCAAGGC
*procA*3.3_NotI_RP	ATAATTTAGCGGCCGCCTATGCGCGGC

3.4.3 Generation of dehydrated ProcA3.3

A sample of 1 mL of ProcA3.3 (6 mg/mL) was oxidized with 5 mM oxidized glutathione (GSSG) at room temperature for 3 h. The peptide was purified using a C4 SPE column (Vydac), and the eluted peptide was lyophilized. The oxidized peptide was dissolved in 1 mL of ddH_2O. An in vitro dehydration reaction was set up with 50 mM HEPES (pH 8.0), 5 mM ATP, 10 mM $MgCl_2$, 0.5 mM GSSG, 50 μM oxidized peptide, and 5 μM ProcM, and incubated at room temperature for 24 h. ProcM was removed from the assay by using Amicon 100 KD filters. The dehydrated peptide was reduced by TCEP and then purified using a C4 SPE column.

3.4.4 PAR assay

All flasks and beakers used were prerinsed with 20% HNO_3 and Milli-Q water. Dialysis was performed with Slidalyzers (Pierce) presoaked in 20 mM HEPES, pH 8.0, containing 10 g/liter Chelex-100 (Bio-Rad). ProcM and ProcM mutants were dialyzed against four changes of 20 mM HEPES, 1 M NaCl, pH 8.0, with 25–50 g/liter Chelex-100 in the beaker, over a total time of 48 h at 4 °C. The Chelex-100-treated protein was then concentrated with Amicon 50 KD

filters that had been presoaked in 20 mM Tris, pH 7.5, and containing 10 g/liter Chelex-100. The reactions were performed in 20 mM HEPES, 1 M NaCl, pH 8.0, that was treated with 25–50 g/liter Chelex-100 and filtered. Aliquots of 40 μL protein solutions or the zinc standard solutions were incubated with 160 μL 100 μM PAR in 4 M guanidine hydrochloride at room temperature. An aliquot of 10 μL 3.5 mM DTNB was then added to the reaction mixture to further release the zinc ion. The absorption of PAR-Zn^{2+} complex at 500 nm was measured, and the concentration of zinc was determined by using a standard curve that was prepared with zinc chloride solutions.

3.4.5 Experimental procedures listed in other chapters

Expression and purification of unmodified and modified peptides (section 2.4.5); Expression and purification of LanM enzymes (section 2.4.6);

In vitro enzymatic assay and endoprotease digestion (section 2.4.7);

Iodoacetamide assay (section 2.4.9);

ESI–LC-MS and tandem MS (section 2.4.10).

3.4.6 Calculated and observed masses in the tandem mass spectra

Table 3.3. Calculated and observed masses in the mass spectra.

***** monoisotopic masses plus one proton were shown for all peaks

Figure	Peak	Calculated mass	Observed mass
Figure 3.4	M- 4 H_2O (ProcA1.1-G-1E)	1783.8	1784.2
	M- 4 H_2O + 1 IAA (ProcA1.1-G-1E)	1840.9	1841.3
	M- 2 H_2O (ProcA2.8)	3435.4	3435.6
	M- 2 H_2O + 1 IAA (ProcA2.8)	3492.4	3492.7
	M- 3 H_2O (ProcA3.3)	2199.0	2198.5
	M- 3 H_2O + 1 IAA (ProcA3.3)	2256.0	2255.2
	M- 5 H_2O (ProcA2.11)	4090.8	4091.4
	M- 5 H_2O + 1 IAA (ProcA2.11)	4147.8	4148.4
	M- 3 H_2O (ProcA3.2)	3222.5	3221.1
	M- 3 H_2O + 1 IAA (ProcA3.2)	3279.5	3278.1
Figure 3.7	M- 2 H_2O (WT)	5035.2	5034.7
	M- 2 H_2O (C924A)		5035.4
	M- 2 H_2O (C970A)		5034.5
	M- 2 H_2O (C971A)		5035.1
	M- 2 H_2O (C971H)		5035.0
Figure 3.8	M- 2 H_2O (WT)	5035.2	5034.7
	M- 2 H_2O (C924A)		5034.8
	M- 2 H_2O (C970A)		5035.0
	M- 2 H_2O (C971A)		5035.2
	M- 2 H_2O (C971H)		5035.4
	M- 2 H_2O + 1 IAA (WT)	5092.2	5091.7
	M- 2 H_2O + 1 IAA(C924A)		5091.8
	M- 2 H_2O + 1 IAA (C970A)		5092.1
	M- 2 H_2O + 1 IAA (C971A)		5092.0
	M- 2 H_2O + 1 IAA (C971H)		5092.4

(Table 3.3 continued)

Figure	Peak	Calculated mass	Observed mass
Figure 3.11	M- 2 H_2O (WT, 15 min)	2125.0	2124.6
	M- 2 H_2O (WT, 30 min)		2124.7
	M- 2 H_2O (WT, 60 min)		2124.5
	M- 2 H_2O (WT, 180 min)		2124.7
	M- 2 H_2O (C971H, 60 min)		2124.6
	M- 2 H_2O (C971H, 180 min)		2124.7
	M- 2 H_2O + 1 IAA (WT, 15 min)	2182.0	2181.5
	M- 2 H_2O + 1 IAA (WT, 30 min)		2181.7
	M- 2 H_2O + 1 IAA (WT, 60 min)		2181.5
	M- 2 H_2O + 1 IAA (WT, 180 min)		2181.7
	M- 2 H_2O + 1 IAA (C971H, 15 min)		2181.6
	M- 2 H_2O + 1 IAA (C971H, 30 min)		2181.7
	M- 2 H_2O + 1 IAA (C971H, 60 min)		2181.6
	M- 2 H_2O + 1 IAA(C971H, 180 min)		2181.7
	M + 2 IAA (WT, 15 min)	2275.1	2274.5
	M + 2 IAA (WT, 30 min)		2274.6
	M + 2 IAA (WT, 60 min)		2274.4
	M + 2 IAA (C971H, 15 min)		2274.5
	M + 2 IAA (C971H, 30 min)		2274.7
	M + 2 IAA (C971H, 60 min)		2274.6

Table 3.4. Calculated and observed masses of each ion resulting from fragmentation in the core region in the tandem MS analysis of the 2-fold dehydrated core peptide ion of ProcA2.8 after digestion by AspN (Figure 3.10). The presence of the ions highlighted in yellow are indicative of the absence of a Dha9-Cys3 ring.

ion	calculated mass	observed peptide **a**	observed peptide **b**
y''8	952.35	952.38	952.43
y''9	1049.40	1049.44	1049.49
y''12	1346.52		1346.58
b36	3690.67		3690.85
b39	3987.79	3988.10	3988.12

Table 3.5. Calculated and observed masses of each ion resulting from fragmentation in the core region in the tandem MS analysis of the 2-fold dehydrated core peptide ion of ProcA1.1 after digestion by GluC (Figure 3.13). The presence of the ions highlighted in yellow are indicative of the absence of a Dhb7-Cys3 ring.

ion	calculated mass	observed peptide **c**	observed peptide **d**
b6	490.25	490.33	490.33
y''5	531.26	531.39	531.39
b7	637.32		637.41
y''6	678.33		1282.66
b8	740.33		740.48
y''7	834.43		834.65
b9	839.40		839.60
y''8	948.47	948.69	948.70
y''9	1019.51	1019.76	1019.76
y''10	1102.55		1102.69
y''11	1159.57		1159.77
y''14	1489.70	1490.08	1490.08
y''15	1636.78	1637.14	1637.10

Table 3.6. Calculated and observed masses of each ion in the tandem MS analysis of the 3-fold dehydrated core peptide ion of ProcA3.3 after digestion by AspN (Figure 3.13 and 3.15). The presence of the ions highlighted in green are indicative of the absence of a Dhb18-Cys21 ring, the ions highlighted in yellow are indicative of the absence of the Dhb18-Cys14 ring, and the ions highlighted in blue are indicative of the absence of the Dhb11-Cys14 ring. Finally, all colored ions are consistent with the absence of a Dhb11-Cys21 ring.

ion	calculated mass	observed peptide **1**	observed peptide **2**	observed peptide **3**	observed peptide **4**	observed peptide **5**	observed peptide **6**
b1	116.03	116.07	116.05			116.09	
b2	199.07	199.13	199.09	199.13	199.13	199.13	199.13
y''2	246.16	246.23	246.19				
b3	256.09	256.23	256.12	256.17	256.17	256.17	256.17
y''4	480.21			480.34			480.37
y''5	608.30			608.47			608.49
y''7	748.36						748.56
y''8	805.38		805.46			805.59	805.66
y''9	968.44						968.71
y''11	1128.48			1128.76		1128.57	1128.80
y''12	1199.51			1199.85		1200.09	1199.85
y''13	1282.55	1282.92	1282.66	1282.90	1282.90	1282.91	1282.92
y''14	1419.61	1420.01	1419.73	1420.00	1420.01	1420.01	1420.08
y''15	1532.69	1533.13	1532.83	1533.11		1533.12	1533.12
y''16	1631.76	1632.27	1631.88	1632.15	1632.15	1632.18	1631.50
y''17	1702.80	1703.29	1702.93	1703.24	1703.26	1703.27	1703.30
y''18	1830.86	1831.35	1831.03	1831.30	1831.45	1831.38	1831.35
y''19	1943.94	1944.47	1944.08		1944.42	1944.45	1944.39
y''20	2000.96	2001.46	2001.09	2001.48	2001.47	2001.49	2001.46
y''21	2084.00	2084.53	2084.13	2084.55	2084.55	2084.60	2084.53

3.5 REFERENCES

1. Li, B., Sher, D., Kelly, L., Shi, Y., Huang, K., Knerr, P. J., Joewono, I., Rusch, D., Chisholm, S. W., and van der Donk, W. A. (2010) Catalytic promiscuity in the biosynthesis of cyclic peptide secondary metabolites in planktonic marine cyanobacteria, *Proc. Natl. Acad. Sci. U.S.A. 107*, 10430-10435.

2. Zhang, Q., Yang, X., Wang, H., and van der Donk, W. A. (2014) High divergence of the precursor peptides in combinatorial lanthipeptide biosynthesis, *ACS Chem. Biol. 9*, 2686-2694.

3. Yu, Y., Zhang, Q., and van der Donk, W. A. (2013) Insights into the evolution of lanthipeptide biosynthesis, *Protein Sci. 22*, 1478-1489.

4. Mukherjee, S., and van der Donk, W. A. (2014) Mechanistic Studies on the Substrate-Tolerant Lanthipeptide Synthetase ProcM, *J. Am. Chem. Soc. 136*, 10450-10459.

5. Thibodeaux, C. J., Ha, T., and van der Donk, W. A. (2014) A Price To Pay for Relaxed Substrate Specificity: A Comparative Kinetic Analysis of the Class II Lanthipeptide Synthetases ProcM and HalM2, *J. Am. Chem. Soc. 136*, 17513-17529.

6. Tang, W., and van der Donk, W. A. (2012) Structural characterization of four prochlorosins: a novel class of lantipeptides produced by planktonic marine cyanobacteria, *Biochemistry 51*, 4271-4279.

7. Li, B., and van der Donk, W. A. (2007) Identification of essential catalytic residues of the cyclase NisC involved in the biosynthesis of nisin, *J. Biol. Chem. 282*, 21169-21175.

8. Li, B., Yu, J. P., Brunzelle, J. S., Moll, G. N., van der Donk, W. A., and Nair, S. K. (2006) Structure and mechanism of the lantibiotic cyclase involved in nisin biosynthesis, *Science 311*, 1464-1467.

9. Paul, M., Patton, G. C., and van der Donk, W. A. (2007) Mutants of the zinc ligands of lacticin 481 synthetase retain dehydration activity but have impaired cyclization activity, *Biochemistry 46*, 6268-6276.

10. Zhang, Q., Yu, Y., Velásquez, J. E., and van der Donk, W. A. (2012) Evolution of lanthipeptide synthetases, *Proc. Natl. Acad. Sci. U. S. A. 109*, 18361-18366.

11. Harris, C. M., Derdowski, A. M., and Poulter, C. D. (2002) Modulation of the zinc(II) center in protein farnesyltransferase by mutagenesis of the zinc(II) ligands, *Biochemistry 41*, 10554-10562.

12. Penner-Hahn, J. (2007) Zinc-promoted alkyl transfer: a new role for zinc, *Curr. Opin. Chem. Biol. 11*, 166-171.

13. Wilker, J. J., and Lippard, S. J. (1997) Alkyl transfer to metal thiolates: kinetics, active species identification, and relevance to the DNA methyl phosphotriester repair center of *Escherichia coli* Ada, *Inorg. Chem. 36*, 969-978.

14. Morlok, M. M., Janak, K. E., Zhu, G., Quarless, D. A., and Parkin, G. (2005) Intramolecular N-H...S hydrogen bonding in the zinc thiolate complex [Tm(Ph)]ZnSCH2C(O)NHPh: a mechanistic investigation of thiolate alkylation as probed by kinetics studies and by kinetic isotope effects, *J. Am. Chem. Soc. 127*, 14039-14050.

15. Hightower, K. E., and Fierke, C. A. (1999) Zinc-catalyzed sulfur alkylation: insights from protein farnesyltransferase, *Curr. Opin. Chem. Biol. 3*, 176-181.

16. Yang, X., and van der Donk, W. A. (2015) The Michael-type cyclizations in lantibiotic biosynthesis are reversible, *ACS Chem. Biol.*, DOI: 10.1021/acschembio.1025b00007.

17. Xie, L., and van der Donk, W. A. (2004) Post-Translational Modifications during Lantibiotic Biosynthesis, *Curr. Opin. Chem. Biol. 8*, 498-507.

18. Hunt, J. B., Neece, S. H., and Ginsburg, A. (1985) The use of 4-(2-pyridylazo)resorcinol in studies of zinc release from Escherichia coli aspartate transcarbamoylase, *Anal. Biochem. 146*, 150-157.

19. Okeley, N. M., Paul, M., Stasser, J. P., Blackburn, N., and van der Donk, W. A. (2003) SpaC and NisC, the cyclases involved in subtilin and nisin biosynthesis, are zinc proteins, *Biochemistry 42*, 13613-13624.

20. Lubelski, J., Khusainov, R., and Kuipers, O. P. (2009) Directionality and Coordination of Dehydration and Ring Formation during Biosynthesis of the Lantibiotic Nisin, *J. Biol. Chem. 284*, 25962-25972.

21. Lee, M. V., Ihnken, L. A., You, Y. O., McClerren, A. L., van der Donk, W. A., and Kelleher, N. L. (2009) Distributive and directional behavior of lantibiotic synthetases revealed by high-resolution tandem mass spectrometry, *J. Am. Chem. Soc. 131*, 12258-12264.

22. Jungmann, N. A., Krawczyk, B., Tietzmann, M., Ensle, P., and Süssmuth, R. D. (2014) Dissecting Reactions of Nonlinear Precursor Peptide Processing of the Class III Lanthipeptide Curvopeptin, *J. Am. Chem. Soc. 136*, 15222-15228.

23. Picot, D., Ohanessian, G., and Frison, G. (2008) The alkylation mechanism of zinc-bound thiolates depends upon the zinc ligands, *Inorg. Chem. 47*, 8167-8178.

24. Rombach, M., Seebacher, J., Ji, M., Zhang, G., He, G., Ibrahim, M. M., Benkmil, B., and Vahrenkamp, H. (2006) Thiolate alkylation in tripod zinc complexes: a comparative kinetic study, *Inorg. Chem. 45*, 4571-4575.

25. Shi, Y., Yang, X., Garg, N., and van der Donk, W. A. (2011) Production of lantipeptides in *Escherichia coli*, *J. Am. Chem. Soc. 133*, 2338-2341.

CHAPTER IV: STUDYING THE INDIVIDUAL DOMAINS OF CLASS II LANTHIPEPTIDE SYNTHETASES USING TRUNCATED ENZYMES

4.1 INTRODUCTION

As discussed in chapter I, the four classes of lanthipeptides are modified by different lanthionine synthetases (Figure 1.4). Class II lanthipeptide synthetases are typically 900-1,200 residues in length and are composed of a proposed dehydratase domain at the N-terminus and a cyclase domain at the C-terminus (Figure 1.4).[1-3] The LanM dehydratase domains have no homology to class I dehydratase LanBs or other characterized proteins in the databases, while the cyclase domains have ~25% sequence identity to LanC, including the conserved residues that have been reported to be crucial for cyclization in the nisin cyclase NisC (Figure 3.2).[4] Similar to LanBs and LanCs, LanM enzymes do not share high degree of sequence identity with each other.[1]

The first successful *in vitro* reconstitution of a LanM enzyme was reported in 2004 by Xie and co-workers, using the precursor peptide LctA and the enzyme LctM generated by heterologous expression in *E. coli*.[5] LctM is the lanthionine synthetase for lacticin 481, a tricyclic lantibiotic isolated from *Lactococcus lactis* subsp. *lactis*. CNRZ 481 in the early 1990s.[6] The proposed biosynthetic pathway for lacticin 481 is shown in Figure 4.1. This report provided a platform for mutagenesis studies of LctM[7, 8] and the production of lacticin 481 analogs including incorporation of nonproteinogenic amino acids, which in several cases improved the antimicrobial activity.[9, 10, 11] LctM utilizes ATP and Mg^{2+} to phosphorylate serine/threonine residues in the LctA precursor peptide and to subsequently eliminate the

resulting phosphate ester to yield Dha/Dhb (Figure 1.1A). LctM does not contain any known ATP-binding domain at the sequence level, but conserved residues involved in phosphorylation resemble critical catalytic residues in serine/threonine kinases.[8] The cyclization on the other hand is ATP-independent and Zn^{2+}-dependent (Figure 1.1B).[7]

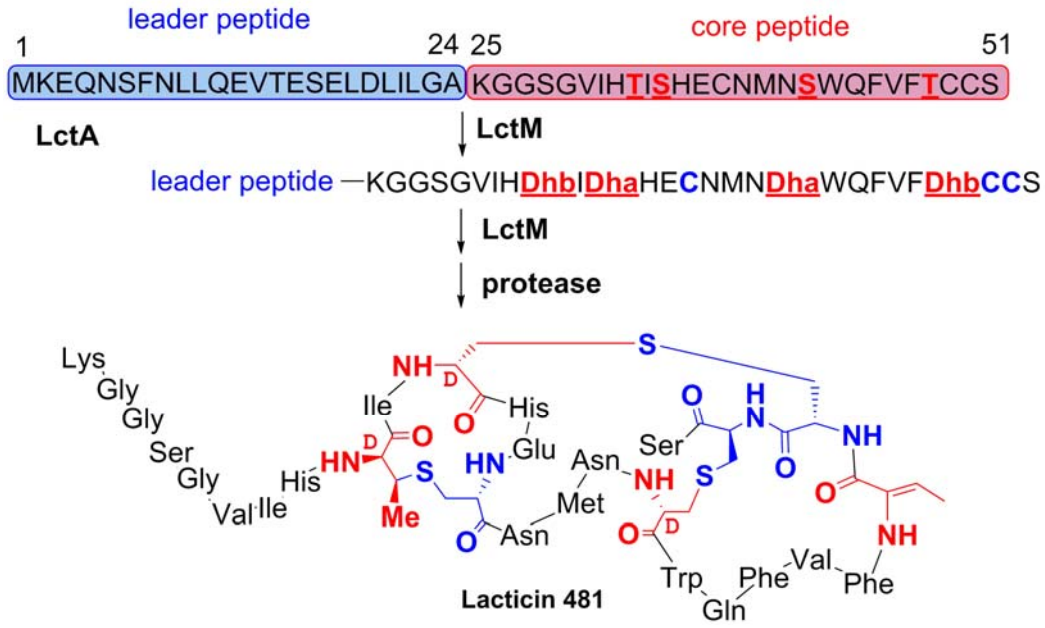

Figure 4.1. The posttranslational maturation process of lacticin 481.[5] The precursor peptide LctA is synthesized ribosomally, followed by dehydration of Ser and Thr (red) residues in the core peptide region of LctA and the conjugate addition of three Cys residues in a regio- and stereospecific manner to three of the Dha and Dhb (red) residues to generate three cyclic thioethers. The leader peptide is proteolytically removed by the N-terminal protease domain of the LctT ABC-type transporter that excretes the final product. Abu, 2-aminobutyric acid.

Lacticin 481 demonstrates inhibitory activity against *L. lactis.* subsp. *cremoris* HP with an MIC value of 1.56 μM,[10] which is only moderate compared to more potent lantibiotics such as nisin (MIC = 32 nM).[12] A variety of other lantibiotics contain a similar ring topology as lacticin 481, including nukacin ISK-1 and members of the salivaricin family. These lantibiotics are often referred to as the lacticin 481 group, and they display diverse antimicrobial scope and potency.[13] The A ring of the lacticin 481 group is similar to the C ring of mersacidin, which is a known lipid II binding motif.[14] A study in 2009 demonstrated lacticin 481 may form complexes with lipid II using thin-layer chromatography.[15] In 2012, another study demonstrated that the A ring of nukacin ISK-1 is the binding motif for lipid II using isothermal titration calorimetry.[16] Together, these reports suggest that the lacticin 481 group demonstrate antimicrobial activities against gram-positive bacteria through the interaction with the membrane component lipid II.

The enterococcal cytolysin is a two-component lantibiotics produced by many clinical isolates of *Enterococcus faecalis*.[17-19] Cytolysin is made up of two separate lanthipeptides $CylL_L$" and $CylL_S$" and is toxic to a broad range of gram-positive bacteria.[18] Besides its antimicrobial activities, cytolysin also possesses unique hemolytic activity against eukaryotic cells, which contributes to the virulence of the pathogenic strains.[20, 21] The linear precursor peptides for both components of cytolysin ($CylL_L$ and $CylL_S$) are modified by the same lanthionine synthetase CylM to install the Lan/MeLan rings,[22] and these cyclized peptides are then digested to remove most of the leader peptide and secreted to the extracellular space by CylB.[23] Next, another protease CylA removes the six additional residues from the leader peptide to generate the mature products. In a recent work published by Tang and co-workers,[24]

the structures of $CylL_L$" and $CylL_S$" were for reported (Figure 4.2A). Interestingly, this study for the first time reported lanthipeptides with other stereochemistry of the Lan/MeLan structures besides the canonical DL-configuration (Figure 4.2B).[24] The Dhx-Dhx-Xxx-Xxx-Cys motif in the precursor peptides have been proven to be important for the unusual LL-configuration,[25] but the role of CylM in the regioselective ring formation has not been fully investigated.

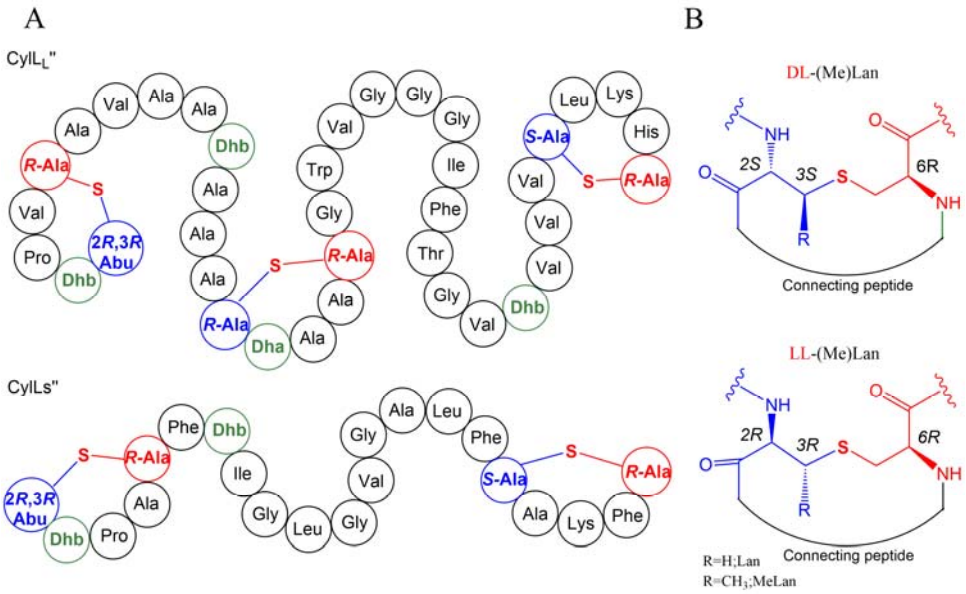

Figure 4.2. A. Structures of the large and small subunits of cytolysin from *E. faecalis.* **B.** The observed stereochemistry of the different rings. The stereochemistry of the α-carbon originating from Cys is *R* in both cases, but the stereochemistry at the α and β carbons originating from Thr (R = Me) is inverted showing that the two types of rings are formed by *anti* addition to opposite faces of the dehydrobutyrine (or dehydroalanine for R = H).

As discussed in chapters II and III, prochlorosins were discovered in 2010 from the planktonic marine cyanobacterium *Prochlorococcus* MIT9313 in a genome-mining effort to search for novel lanthipeptides in the unexplored genomes.[26] ProcM and its close homologues with three zinc-binding cysteines form a distinct clade in the phylogenetic tree of LanM (Figure 1.5),[27] and generally (but not always) are associated with multiple substrate peptides with diverse sequences.[28]

For all the class II lanthipeptide synthetases, no crystal structures have been reported so far. In fact, before 2014, the only crystal structure reported for any lanthionine synthetase was for the nisin cyclase NisC.[29] In 2014, the dehydratase for nisin biosynthesis (NisB) was also successfully crystallized to provide detailed structure information for deeper understanding of class I lanthipeptide biosynthesis. A possible reason of the lack of structural information of LanM enzymes is that they are multidomain proteins and relatively large in size. As the cyclase domains of LanMs are homologous to LanCs, it is reasonable to speculate that the individual cyclase domain might have similar behaviors as LanCs and might crystallize under the similar conditions used for NisC crystallization. The three LanMs discussed above are responsible for the production of distinct products: LctM processes a single product with Lan/MeLan rings of DL-configuration; CylM generates two products with a mixture of DL- and LL-Lan/MeLan rings; while ProcM manages to modify 30 products with DL-Lan/MeLan rings. While it is difficult to obtain information about the difference of these enzymes at the sequence level, the structures of these enzymes might reveal the unique properties of each enzyme. The idea of truncating LanM into individual domains was initially motivated by the possibility to obtain crystal structures of these domains to allow further understanding of class II lanthipeptide

biosynthesis. Also, obtaining individual domains of LanMs for the first time allows studying the dehydration and cyclization of class II lanthipeptide in isolation, which is important to understand the product selectivities of LanM enzymes. Moreover, this study will also provide insights on the evolution of the lanthionine synthetases as it is proposed that the LanM enzymes might have evolved from a fusion event of two ancestor enzymes that are responsible for dehydration and cyclization respectively.

4.2 RESULTS

4.2.1 Cloning, heterologous expression, purification and crystallization attempts of truncated LanM enzymes

To determine the possible cleavage sites to split LanMs into two individual domains, the amino acid sequences of several LanMs were aligned with NisC using the NCBI online alignment tool Cobalt.[30] As indicated in Figure 4.3, pRSF-Duetl constructs containing the following truncated enzymes were made by molecular cloning methods as described in the Experimental Procedures section 4.4.2: ProcM-1-660, ProcM-655-1068, LctM-1-543, LctM-544-922, CylM-1-658, and CylM-659-993.

These truncated enzymes were expressed heterologously in *Escherichia coli* as N-terminally His_6-tagged fusion proteins. Expression of the LanM proteins using the pRSF-Duet1 vector resulted in higher expression level as well as less protein produced in the inclusion bodies than using the pET28b vector. The purification processes are described in detail in section 4.4.3. ProcM-655-1068, LctM-1-543 and CylM-1-658 were expressed solubly, and the proteins were isolated as monomers and used for downstream experiments. ProcM-1-660 was

also expressed solubly, but existed in various oligomeric states. The oligomers of lower molecular mass were collected and used for downstream experiments. LctM-544-922 was expressed at a reasonable level, but protein of the predicted mass was only found in the insoluble portion of the cell lysate. CylM-659-993 was handed to group member Weixin Tang for further characterization. ProcM-1-660, ProcM-655-1068 and LctM-1-543 were used to set up multiple 96-well trays to test different crystallization conditions, but no crystal was observed in all experiments. CylM-1-543 was also used for crystallization attempts by Dr. Shihui Dong in Professor Satish Nair Lab, but also no crystal was obtained although full length CylM was successfully crystallized (unpublished data when this chapter is written).

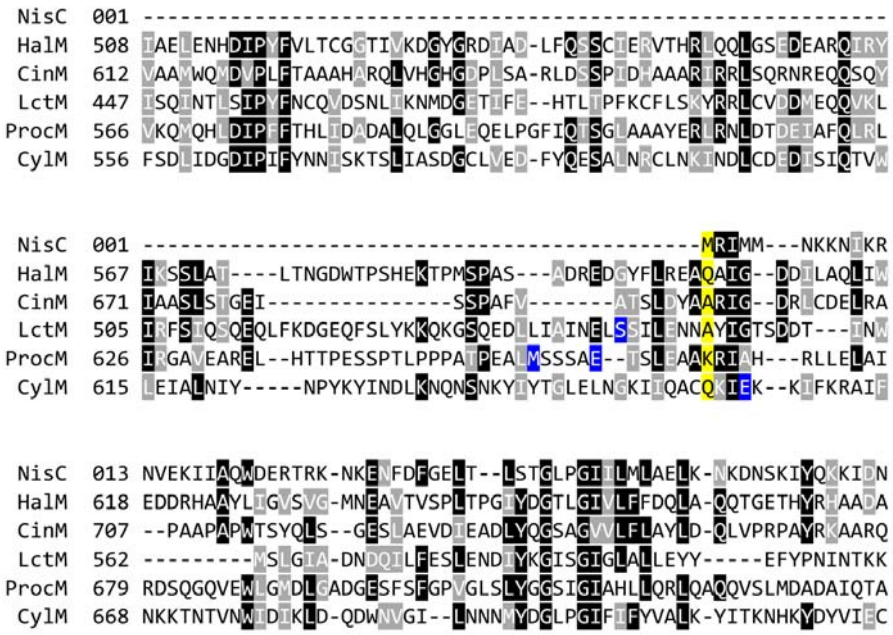

Figure 4.3. Sequence alignment of several LanMs and NisC showing the possible boundary between the dehydration and cyclization domains. The residues highlighted in yellow are the residues at the same position as the first residue of NisC. The residues highlighted in blue indicate the sites for the truncation, which were chosen to be within a few amino acids from the highlighted residues.

4.2.2 The dehydratase domain of ProcM modifies its substrates with reduced rate

ProcM-1-660 includes the putative dehydratase domain. The activity of ProcM-1-660 was tested in parallel with WT-ProcM against two ProcA substrate peptides: ProcA1.1 and ProcA3.3. The *in vitro* enzymatic assays were carried out at room temperature using the standard conditions for ProcM: 10 mM ATP, 1 mM TCEP, 50 mM HEPES (pH 8.0), and various amounts of peptide and enzyme. As shown in Figure 4.4A and Figure 4.4B, ProcM-1-660 was able to catalyze one dehydration of both ProcA1.1 and ProcA3.3. Even though the majority of the starting peptides were converted to the one-fold dehydrated species, no signals corresponding to the two-fold dehydrated peptide were detected. It is obvious that the dehydration activity of ProcM-1-600 was significantly reduced compared to WT-ProcM, but it is still interesting to see the dehydratase domain alone was active. To test if the specificity of the enzyme was retained, ProcM-1-660 was also tested against LctA. As shown in Figure 4.4C, no dehydration was observed in LctA. Since ProcM was able to modify a chimeric peptide consisting of the ProcA leader peptide and LctA core peptide,[27] the dehydratase domain of ProcM must specifically recognize the ProcA leader peptide in the chimeric peptide analog.

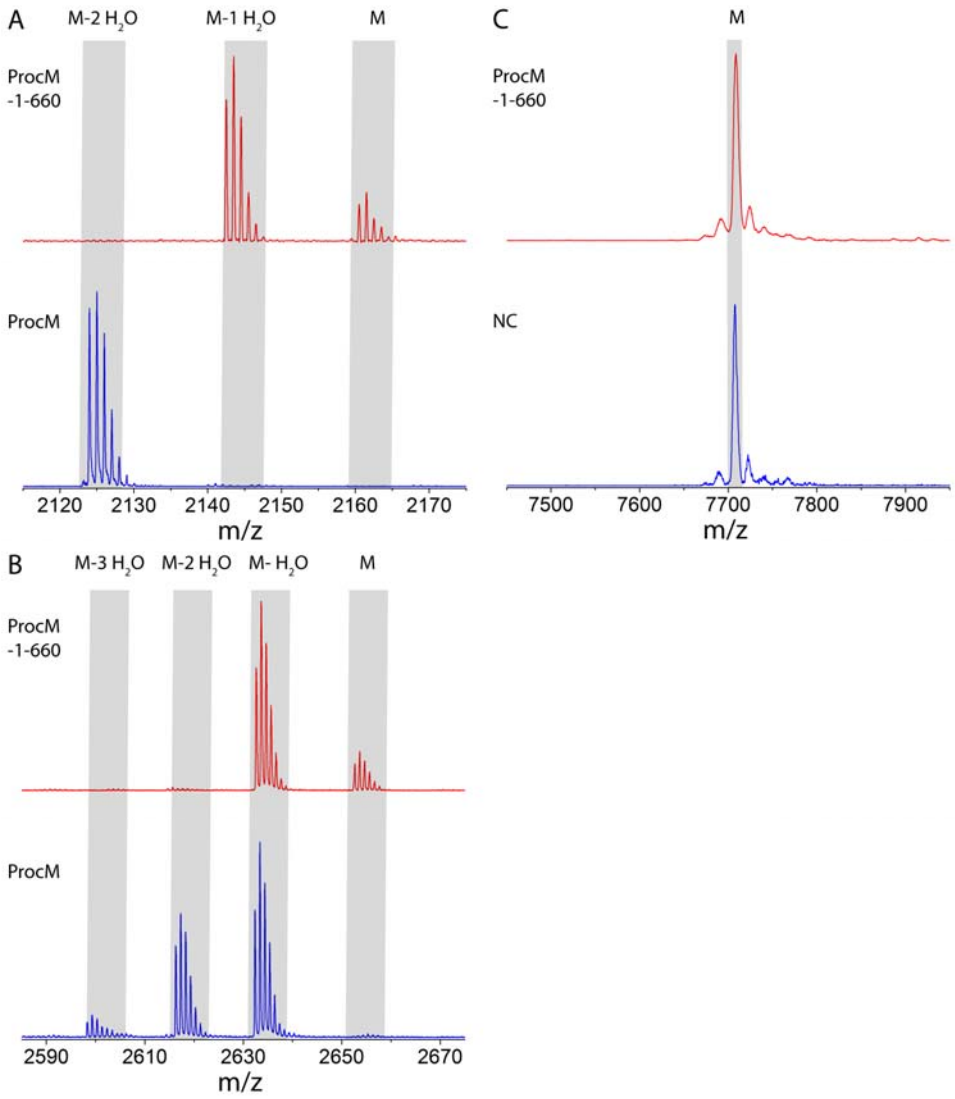

Figure 4.4. MALDI-TOF mass spectra of (**A**) ProcA1.1 peptide (25 μM) modified by either 2 μM WT ProcM (blue) or ProcM-1-660 (red) for 45 min, followed by subsequent digestion by GluC; **(B)** ProcA3.3 peptide (25 μM) modified by either 2 μM WT ProcM (blue) or ProcM-1-660 (red) for 45 min, followed by subsequent digestion by GluC; (**C**) LctA peptide (2 μM) incubated with either 2 μM ProcM-1-660 (red) or no enzyme (blue) for 3 h. M=unmodified starting peptide.

4.2.3 The cyclase domain of ProcM determines the ring formation of the products

ProcM-655-1068 includes the putative cyclase domain. In order to test the cyclization activity of ProcM-655-1068, the first requirement was the generation of linear dehydrated ProcA. ProcA3.3 was treated with 5 mM oxidized glutathione to introduce a disulfide bond between the two cysteines, which prevents the cysteines from attacking the unsaturated amino acids during the dehydration process. WT ProcM was able to take the peptide with the disulfide bridge as substrate and catalyzed the 3-fold dehydration of the core peptide. The activity of ProcM-655-1068 was tested using this dehydrated ProcA3.3, and the cyclization reaction was monitored by the iodoacetamide (IAA) assay described in section 2.4.9. ProcM-655-1068 was able to fully cyclize ProcA3.3 and the majority of the product did not react with IAA (Figure 4.5). The negative control shows the level of non-enzymatic cyclization in the *in vitro* enzymatic assay, and only about 20% peptide formed one ring non-enzymatically, and no fully cyclized peptide was detected (Figure 4.5). This result clearly demonstrated that the cyclase domain was fully active on its own.

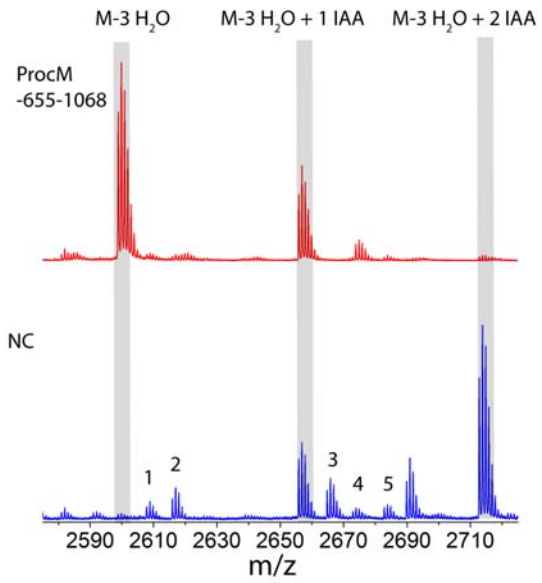

Figure 4.5. MALDI-TOF mass spectra of dehydrated ProcA3.3 peptide (25 μM) incubated with either 2 μM ProcM-655-1068 (red) or reaction buffer with no enzyme (blue) for 3 h, followed by subsequent digestion by GluC and IAA treatment. The minor peaks in the negative control (NC) correspond to incomplete dehydration: 1, M − 2 H_2O; 2, M − H_2O; 3, M − 2 H_2O + 1 IAA; 4, M − H_2O + 1 IAA; 5, M + 1 IAA. These species were present at lower intensity in the top mass spectrum due to the better ionization of the cyclized peptide.

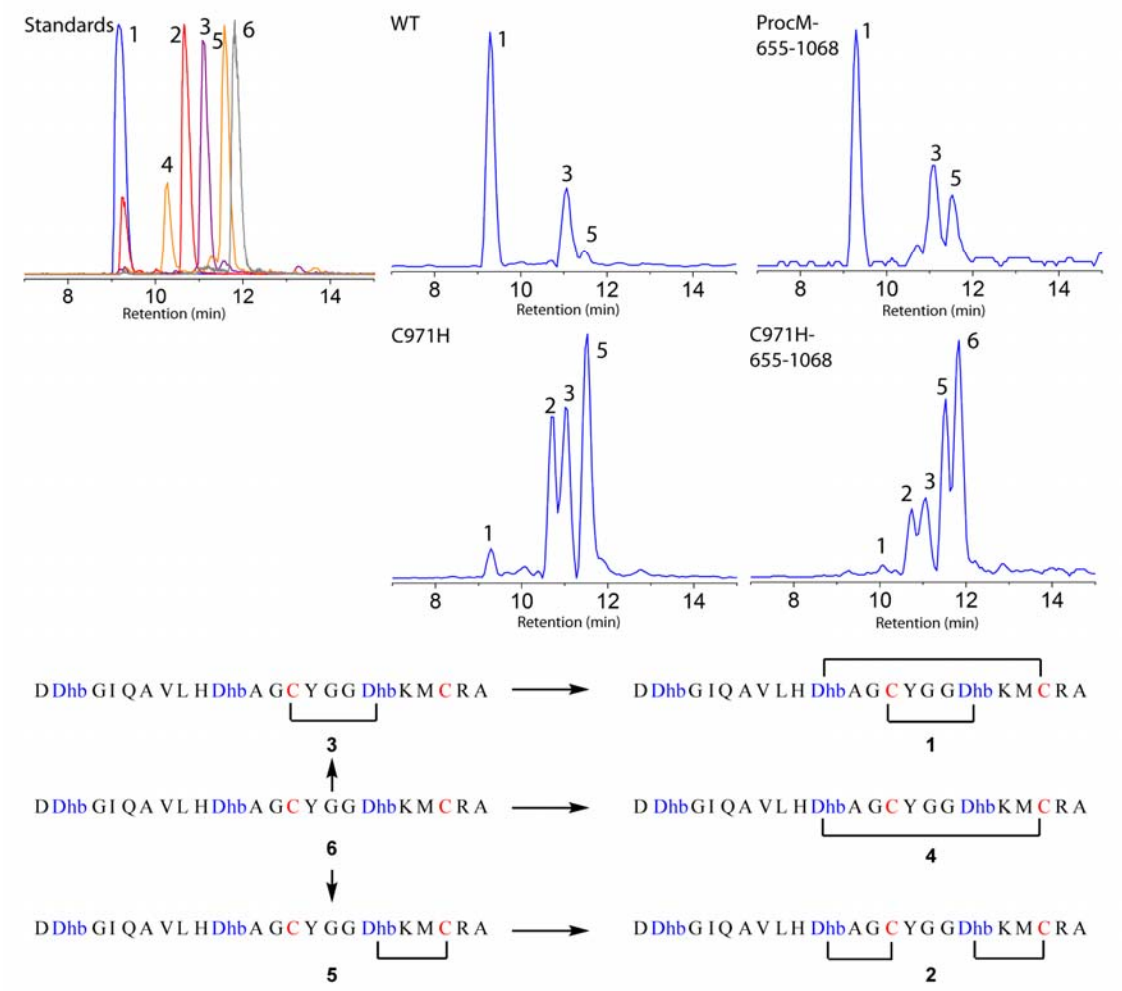

Figure 4.6. EICs (Extracted ion chromatograms) of the ESI-LC-MS experiments showing the Δ1 core peptide ion obtained by endoprotease AspN digestion of dehydrated ProcA3.3 incubated with WT ProcM, ProcM-655-1068, ProcM-C971H, and ProcM-C971H-655-1068. The generation of standards was described in the legend of Figure 3.14. All the peptides were digested by AspN to remove the leader peptide.

As described in chapter III, a single mutation of the zinc ligand Cys971 to a histidine residue altered the regioselectivity of the enzyme and resulted in the formation of a different major product for ProcA3.3. To test if this regioselectivity was retained with ProcM-655-1068,

dehydrated ProcA3.3 was incubated with either the WT-ProcM or ProcM-655-1068 for 3 h, subsequently digested by AspN protease and analyzed by liquid chromatography electrospray ionization mass spectrometry (ESI-LC-MS) to differentiate the ProcA3.3 core peptide with different potential cyclization patterns. As depicted in Figure 4.6, both the full length enzyme and the isolated cyclase domain prefer the formation of product **1** and intermediate **3**. Interestingly, both the full length ProcM-C971H and its isolated cyclase domain with the C971H mutation prefer the formation of product **2** and intermediate **5** instead, which strongly suggests that the cyclase domain of ProcM is able to determine the ring formation of ProcAs.

As shown in section 4.2.2, the dehydration activity of ProcM is leader-dependent, which suggests a leader binding site in the dehydratase domain. To test whether the cyclase domain has its own leader binding site or whether it depends on the dehydratase domain for substrate binding, the leader-dependence of ProcM-655-1068 was tested using dehydrated ProcA3.3 core peptide (peptide generation described in section 4.4.4). In Figure 4.7, it is shown that there was no difference in the IAA assay between the enzyme-treated core peptide and the negative control. This experiment suggests that the activity of the cyclase domain is also leader-dependent, and ProcM may have two independent leader binding sites, one in each of its two domains.

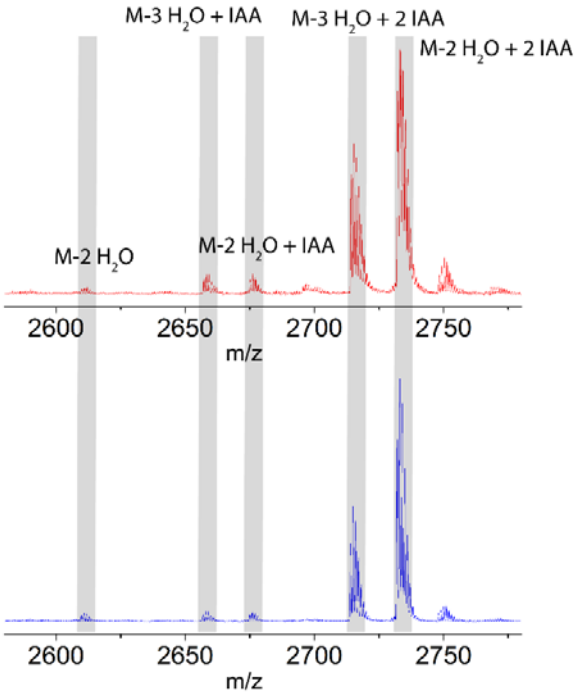

Figure 4.7. MALDI-TOF mass spectra of dehydrated ProcA3.3 core peptide (~20 μM) incubated with 5 μM ProcM-655-1068 (red) or no enzyme (blue) for 3 h, followed by subsequent digestion by GluC and treatment with IAA.

4.2.4 Synergistic action of two ProcM domains

As discussed in section 4.2.2, ProcM dehydratase domain was only able to catalyze one-fold dehydration. This observation suggested the possibility that the two domains collaborate with each other to modify the substrate, and that the cyclization domain might facilitate the dehydration reaction. Indeed, it has been reported for nisin, haloduracin, and curvopeptin that the dehydration and cyclization reactions are alternating processes during biosynthesis.[31-34] Although ProcM catalyzes all dehydration reactions before commencing cyclization processes for the substrates that have been investigated thus far,[33, 35] no studies were carried out with individual dehydratase and cyclase domains. It would be interesting to see if the dehydration and cyclization domains show cooperative behavior. Surprisingly, a dramatic improvement in

the dehydration of both ProcA1.1 and ProcA3.3 was indeed observed when the cyclase domain was added (Figure 4.8). Not only did the ProcM-1-660 catalyze 3-fold dehydration in presence of the cyclase domain, the rate of the dehydration reaction was also increased to a level similar as that of the WT enzyme.

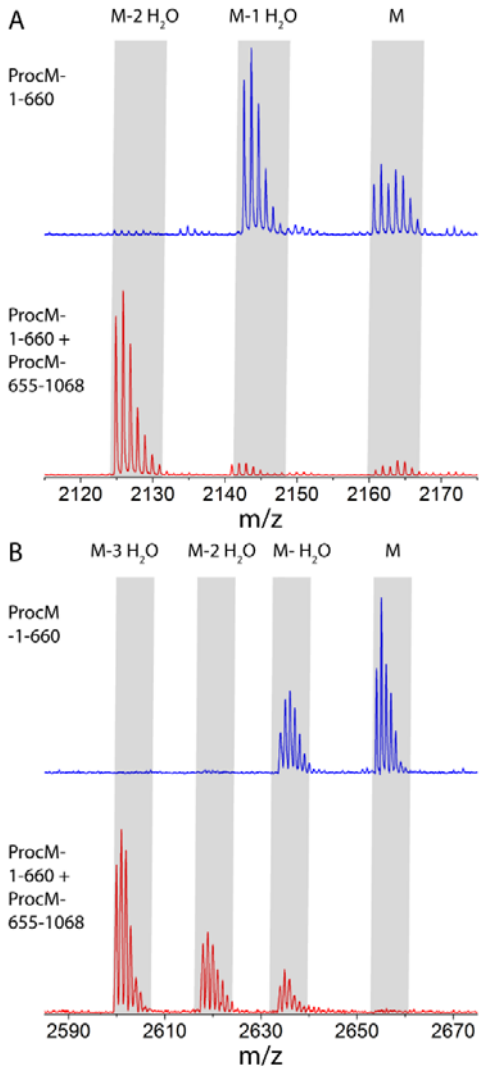

Figure 4.8. MALDI-TOF mass spectra of (**A**) ProcA1.1 (25 μM) modified by 2 μM ProcM-1-660 with (red) or without ProcM-655-1068 (blue) for 1 h, and subsequently digested by GluC; **(B)** ProcA3.3 (25 μM) modified by 2 μM ProcM-1-660 with (red) or without ProcM-655-1068 (blue) for 1 h, followed by subsequent digestion by GluC.

4.2.5 The binding affinity of ProcA3.2 leader peptide to ProcM individual domains

As discussed in section 4.2.2 and 4.2.3, ProcM may have two independent leader binding sites since both the dehydratase and cyclase domains of ProcM process leader peptide-dependent activities. In order to confirm that ProcA leader peptide indeed binds to these two domains, fluorescence polarization experiments were used to evaluate the binding affinity between ProcA leader peptide and truncated ProcM enzymes. His_6-ProcA3.2 leader peptide was obtained using the construct generated as described in section 4.4.5, and fluorescently labelled with fluorescein isothiocyanate (FITC) which primarily reacts with the N-terminal free amine group of the peptide (section 4.4.6). Using procedures described in section 4.4.7, the binding affinity of FITC-ProcA3.2Lea to WT ProcM, ProcM-655-1068, and ProcM-1-660 were measured. As shown in Figure 4.9, the cyclase domain of ProcM showed a binding affinity similar to WT ProcM (8.0 ± 1.2 μM). The binding of FITC-ProcA3.2Lea to the dehydratase domain was also detected, but the binding affinity was about 8-fold weaker compared to that to WT ProcM.

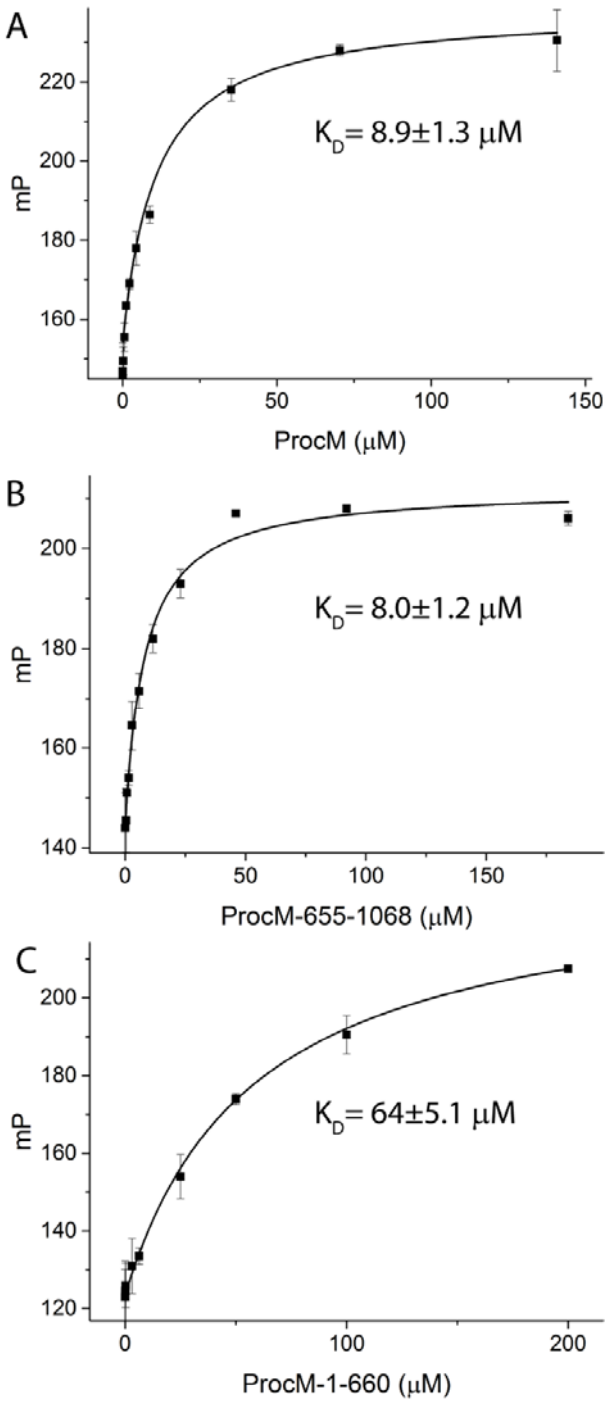

Figure 4.9. Fluorescence polarization assays with WT ProcM (A), ProcM-655-1068 (B) and ProcM-1-660 (C) using 10 nM FITC-ProcA3.2Lea.

4.2.6 The dehydratase domains of LctM and CylM are fully active

Residues 1-543 constitute the putative dehydratase domain of LctM. The activity of LctM-1-543 was tested in parallel with WT-LctM against LctA peptide. Unlike ProcM-1-660, LctM-1-543 was able to catalyze all four possible dehydrations in LctA, and with a similar efficiency compared to the WT enzyme (Figure 4.10). Notably, the major intermediate observed for WT enzyme was two-fold dehydrated LctA, but this intermediate did not accumulate in the reaction catalyzed by LctM-1-543 for unknown reason.

LctM was reported to be able to catalyze the *in trans* modifications of truncated LctA core peptide in the presence of the LctA leader peptide. [10] It was of interest to see if the dehydratase domain could do the same. Indeed, LctM-1-543 was able to modify the LctA core peptide when the leader peptide was provided *in trans*, again with a similar rate as the WT enzyme (Figure 4.11). As will be discussed in chapter V, the binding of LctA leader peptide to LctM-1-543 was confirmed using fluorescence polarization experiments with a K_D value of 9.6±2.1 μM.

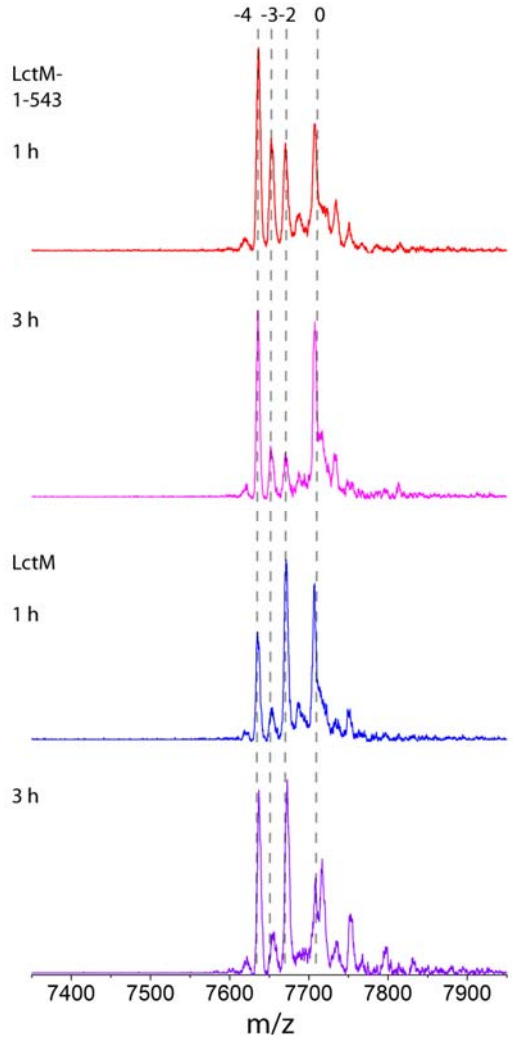

Figure 4.10. MALDI-TOF mass spectra of LctA (20 μM) modified by 5 μM LctM-1-543 or LctM for 1 h, or 3 h. The numbers above the peaks correspond to the number of dehydration events.

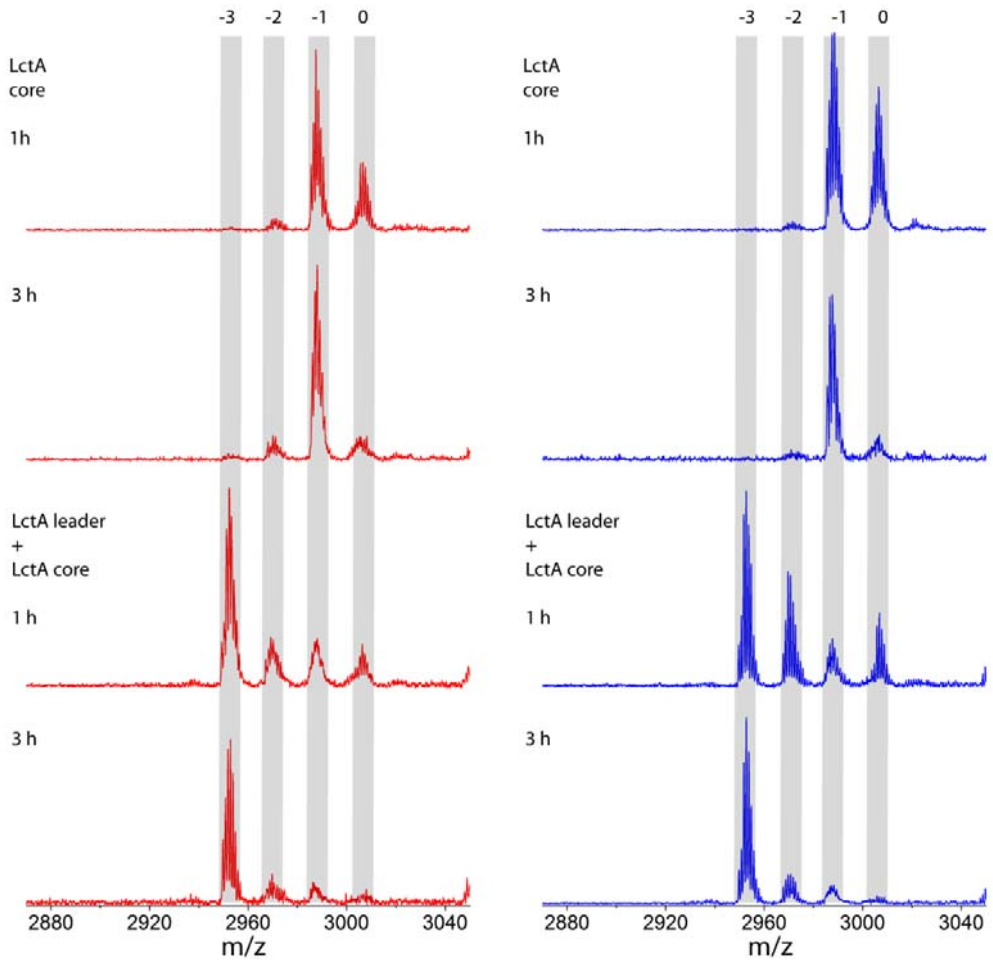

Figure 4.11. MALDI-TOF mass spectra of LctA N15R/F21H core peptide (20 μM) modified by 5 μM LctM-1-543 (red) or LctM (blue) with or without 2 μM LctA leader peptide for 1 h, or 3 h. The numbers above the peaks correspond to the number of dehydration events.

Residues 1-658 form the putative dehydratase domain of CylM. Like the truncated LctM, CylM-1-658 was also able to fully dehydrate one of its substrate (CylLs) (Figure 4.12). Besides the fully dehydrated product (4-fold dehydrated), a peak corresponding to the 5-fold dehydrated product and a peak corresponding to a 4-fold dehydrated and phosphorylated product were also

detected. This observation might indicate one extra dehydration in the leader region because the core peptide contains only four Ser and Thr residues (Figure 4.3).

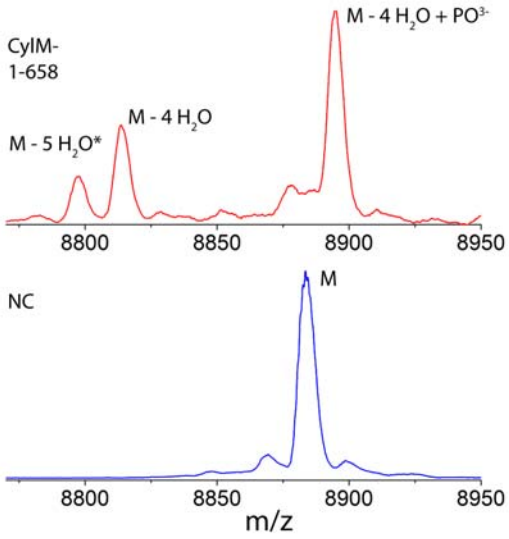

Figure 4.12. MALDI-TOF mass spectra of CylLs (10 μM) incubated with 5 μM CylM-1-548 (red) or no enzyme (blue) for 1 h. The peak labeled with the star symbol corresponds to the five-fold dehydrated peptide. Since the maximum number of dehydrations that can occur in the CylLs core peptide is four, it is possible that the serine at the -7 position in the leader peptide is dehydrated. NC=negative control.

4.2.7 The chimeric enzymes

As mentioned in a previous study,[36] ProcM was able to produce active lacticin 481 using a chimeric peptide containing the ProcA leader peptide and LctA core peptide. However, this strategy did not work for all lanthipeptides as ProcM only catalyzed partial modifications in the NisA and ElxA core peptides attached to the ProcA leader peptide. As shown in section

4.2.3, the cyclase domain itself determines the ring topology of the product. Thus, it was of interest to see if a chimeric ProcM-LctM enzyme with the highly promising ProcM dehydratase domain and a specific LctA cyclase domain would be active against the native or chimeric ProcA/LctA peptides. LctM-ProcM chimeric enzyme consisting of the LctM dehydratase domain and the ProcM cyclase domain as well as a ProcM-NisC enzyme with the ProcM dehydratase domain and the NisC cyclase were constructed. ProcM-LctM and LctM-ProcM were successfully expressed in *E. coli*, but the ProcM-NisC protein did not express. LctM-ProcM showed weak dehydration activity on LctA (3-fold dehydrated peptide was observed), but no activity against ProcALea-LctAStr (a chimeric peptide consisting of the ProcA leader peptide and the LctA core peptide) and ProcA3.3 (Figure 4.13). Only minor cyclization activity was observed upon LctA incubation with LctM-ProcM, which is attributed to non-enzymatic cyclization. Similarly for ProcM-LctM, weak dehydration activity was observed for ProcA3.3, but not for ProcALea-LctAStr and LctA (Figure 4.14). The cyclization reaction was also negligible.

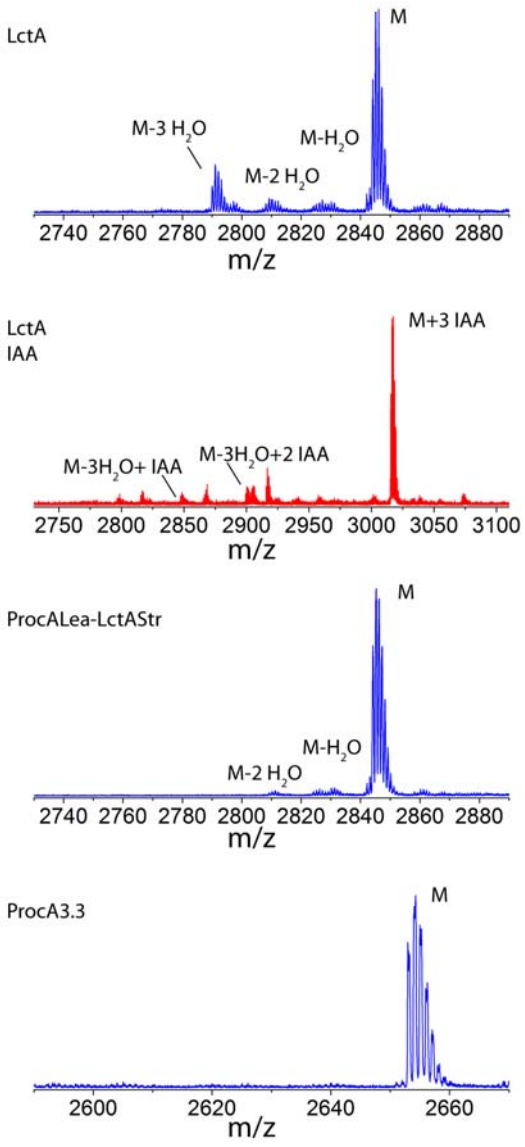

Figure 4.13. MALDI-TOF mass spectra of 25 μM LctA, ProcALea-LctAstr, or ProcA3.3 incubated with 5 μM LctM-ProcM for 3 h, followed by protease digestion and analysis with (red) or without (blue) IAA assay. LctA and ProcALea-LctAstr were digested by LysC, and ProcA3.3 was digested by GluC.

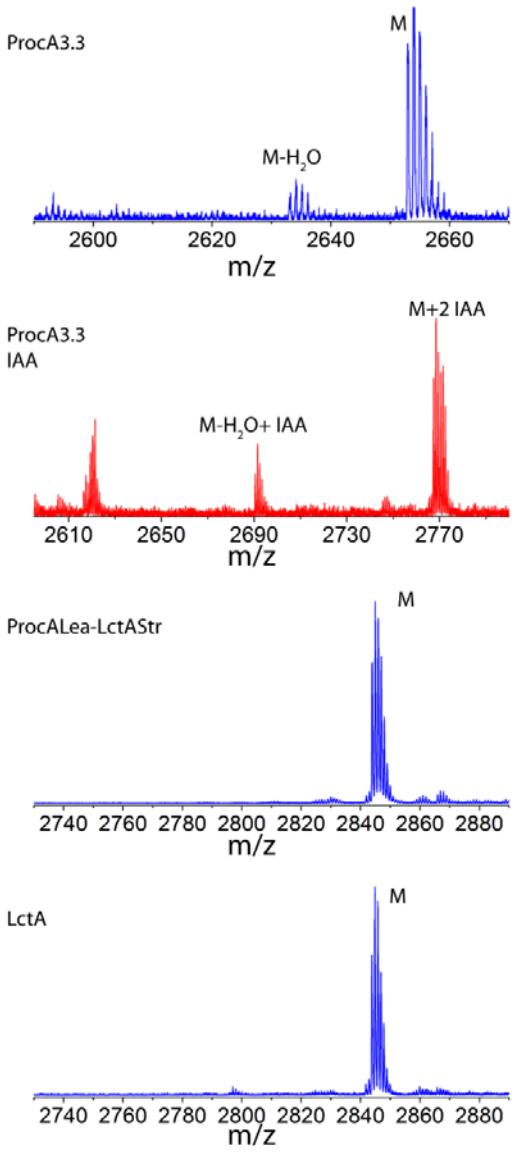

Figure 4.14. MALDI-TOF mass spectra of 25 μM ProcA3.3, ProcALea-LctAstr, or LctA incubated with 5 μM LctM-ProcM for 3 h, followed by protease digestion and analysis with (red) or without (blue) IAA assay. LctA and ProcALea-LctAstr were digested by LysC, and ProcA3.3 was digested by GluC.

4.3 DISCUSSION AND OUTLOOK

In summary of this chapter, the individual dehydration and cyclization domains of three class II lanthionine synthetases (ProcM, LctM, and CylM) were constructed and the proteins were obtained by heterologous expression in *E. coli.* For all the proteins that were solubly expressed, at least partial activities were observed against their native substrate which suggests the two LanM domains are able to function independently. Except for the ProcM dehydratase domain, all other truncated enzyme characterized in this study were able to catalyze full conversion of the substrate, and in the case of LctM dehydratase domain, it was demonstrated to have a similar reaction rate compared to the WT-LctM.

For the dehydratase domain, the truncated enzymes from different LanM displayed different behavior. The LctM and CylM dehydratase domains both showed full activity on their substrates, and these two enzymes are relatively close in a phylogenetic tree (Figure 1.5). Both enzymes belong to the LanM clade from *Firmicutes*, while ProcM and its homologues are in a distinct clade. It is likely that LanM enzymes have several subgroups due to different evolutionary paths, and enzymes from each group might display different catalytic properties. This hypothesis is supported by the fact that the substrate tolerance, reaction rate and regioselectivity of these three LanM are different from each other.

The ProcM dehydratase domain did not have the ability of catalyzing multiple dehydration in the substrate. The reason for this loss of function cannot be explained solely by the decrease of the dehydration rate, since the majority of the starting material was converted to the 1-fold dehydrated species, but no signals corresponding to the 2 or 3-fold dehydrated peptide was observed even with extended incubation time. With the presence of the cyclase domain *in trans*,

the dehydratase domain was able to fully dehydrate ProcAs at a rate that is comparable to the WT-ProcM. Clearly the cyclase domain did not catalyze the dehydration reaction directly, because incubation of the cyclase domain alone with the substrates did not result in dehydration. Thus, there are two possible explanations of this synergistic effect of the two ProcM domains: the first one is that the binding of the cyclase domain induces conformational changes in the dehydratase domain that activates the enzyme; the second one is that cyclization of the partially dehydrated intermediates results in tighter binding to the enzyme and preventing immature substrate release. Because cyclization has been shown to commence only after dehydration is completed,[34, 35] the second explanation does not hold.

The ProcM cyclase domain is the only cyclase tested in this study. Not only did the cyclase domain catalyze full conversion of the dehydrated ProcA3.3 peptide, it also formed product with the correct ring topology. As discussed in chapter III, if a zinc ligand of ProcM was mutated, the mutant enzyme lost ability to convert ProcA3.3 into a product with two-overlapping rings and this change in regioselectivity was accompanied with a decrease in the cyclization rate. It is interesting that the complete removal of the large dehydratase domain seems to bring no changes at all for the cyclase domain. This observation supports the idea that LanM enzymes evolved by a fusion event of a dehydratase domain and the LanC-like ancestor.

Previous work performed by Goto and coworkers demonstrated the phosphorylation activity of the individual kinase domain and the elimination activity of the individual lyase domain of the class IV lanthionine synthetase VenL, but no activity was reported for the individual cyclase domain.[37] Also, in a recent study on NukM,[38] only the dehydratase domain displayed activity on NukA and no activity was reported for the cyclase domain. This chapter

demonstrates a successful example to study the dehydration and cyclization reaction catalyzed by LanM in isolation in order to understand the relationship between these two processes. It is been applied in other studies for the mechanistic investigation of the LanM enzymes,[36] and it can be generally applied for all LanM enzymes. Furthermore, for class III and IV lanthionine synthetases with three domains, studying each domain individually is also beneficial to understand the molecular detains of lanthionine formation and the evolutionary history of these complex enzymes.

4.4 EXPERIMENTAL PROCEDURES

4.4.1 Construction of the expression plasmids for truncated enzymes

The plasmids pET28b-*procM*,[26] pRSFDuet1-*procM*-C971H,[36] pET28b-*lctM*,[5] and pRSFDuet-1-*cylM*-2[24] were isolated from *E. coli* DH5α cells and used as templates for PCR reactions. The DNA sequence for each truncated enzyme was PCR amplified using Phusion polymerase by 30 cycles of denaturing (98 °C for 30 s), annealing (55 °C for 30 s), and extending (72 °C, 1 min for each 1.5 kb) using forward primer and reverse primer indicated in Table 4.1. The PCR product was digested with EcoRI and NotI restriction enzymes and ligated into the first multiple cloning site (MCS1) in the pRSF-Duet1 vector to generate the constructs listed in Table 4.1. The sequences of the inserts were checked by DNA sequencing.

Table 4.1. Primer sequence used in section 4.4.1.

Construct	Primer Name	Primer Sequences (5'-3')
pRSFDuet1-*procM*-1-660 (Template: pET28b-*procM*)	Forward: *procM* _EcoRI_FP	CAGGATCCGAATTCGATGGAA AGTCCATCATCTTGG
	Reverse: *procM* _660_NotI_RP	CTCAGCGCGGCCGCTCACTCA GCCGATGATGACATC
pRSFDuet1-*procM*-655-1068 (Template: pET28b-*procM*)	Forward: *procM*_655_EcoRI_ FP	GAATCCGAATTCGATGTCATCA TCGGCTGAGACGA
	Reverse: *procM* _NotI_RP	AAGGAAAAAAGCGGCCGCTTA TTCAGTAGGCCAGAGA
pRSFDuet1-*procM*-655-1068-C971H (Template: pRSFDuet1-*procM*-C971H)	Forward: *procM*_655_EcoRI_F P	GAATCCGAATTCGATGTCATCA TCGGCTGAGACGA
	Reverse: *procM* _NotI_RP	AAGGAAAAAAGCGGCCGCTTA TTCAGTAGGCCAGAGA
pRSFDuet1-*lctM*-1-543 (Template: pET28b-*lctM*)	Forward: *lctM*_EcoRI_FP	CATATGGAATCCGATGAAAAA AA AGACTTAC
	Reverse: *lctM*_543_NotI_RP	CTCAGCGGCGCGCCTCATGAA AGCTCATTTATCGC
pRSFDuet1-*lctM*-544-922 (Template: pET28b-*lctM*)	Forward: *lctM*_544_EcoRI_FP	GCTGAGGAATTCGAGTATCTT AGAAAACAAT GC
	Reverse: *lctM*_NotI_RP	CTCAGCGCGGCCGCTTAATCAA CATATGGC
pRSFDuet1-*cylM*-1-658 (Template: pRSFDuet-1-*cylM*-2)	Forward: *cylM*_EcoRI_FP	AAAAAGAATTCGGAAGATAAT CTGATTAA T
	Reverse: *cylM_658*_NotI_RP	CTCAGCGCGGCCGCTCATTCAA TTTTCTGGCAGGCC
pRSFDuet1-*cylM*-659-993 (Template: pRSFDuet-1-*cylM*-2)	Forward: *cylM*_659_EcoRI_FP	CAGGATCCGAATTCGAAAAAA ATCTTTAAACGTGC
	Reverse: *cylM*_NotI_RP	AAAAAGCGGCCGCTTACAGTT CAAACAGCA G

4.4.2 Construction of the expression plasmids for chimeric enzymes

The plasmids pET28b-*procM*,[26] pET28b-*lctM*, and pRSFDuet-1-*cylM*-2[24] were isolated from *E. coli* DH5α cells and used as templates for PCR reactions. The DNA sequence for each truncated enzyme was PCR amplified using Phusion polymerase by 30 cycles of denaturing (98 °C for 30 s), annealing (55 °C for 30 s), and extending (72 °C, 1 min for each 1.5 kb) using forward and reverse primers indicated in Table 4.2. The chimeric enzymes were constructed in two steps: 1. the gene for the N-terminal domain (with no stop codon) was inserted between EcoRI and AscI restriction sites in the first multiple cloning site (MCS1) in the pRSF-Duet1 vector. The sequences of the inserts were checked by DNA sequencing. 2. Constructs generated in step 1 were used as vector to insert the gene for the C-terminal domain between AscI and NotI restriction sites. The sequences of the second inserts were checked by DNA sequencing.

Table 4.2. Primer sequence used in section 4.4.2.

Construct	**Primer Name**	**Primer Sequences (5'-3')**
pRSFDuet1-*procM*-1-660-AscI (Template: pET28b-*procM*)	Forward: *procM* _EcoRI_FP	CAGGATCCGAATTCGATGGAAAGTCCATCATCTTGG
	Reverse: *procM*_660_AscI_RP	CTCAGCGGCGCGCCCTCAGCCGATGATGACATC
pRSFDuet1-*procM*-*lctM* (Template: pRSFDuet1-*procM*-1-660-AscI)	Forward: *lctM*_655_AscI_FP	GCTGAGGGCGCGCCAAGTATCTT AGAAAACAAT GC
	Reverse: *lctM* _NotI_RP	CTCAGCGCGGCCGCTTAATCAACATATGGC
pRSFDuet1-*procM*-nisC (Template: pRSFDuet1-*procM*-1-660-AscI)	Forward: *nisC*_AscI_FP	GCTGAGGGCGCGCCAATGAGGATAATGATG
	Reverse: *nisC*_NotI_RP	CTCAGCGCGGCCGCTCATTTCCTCTTCCCTCC
pRSFDuet1-*lctM*-1-543-AscI (Template: pET28b-*lctM*)	Forward: *lctM*_EcoRI_FP	CATATGGAATCCGATGAAAAAAA AGACTTAC
	Reverse: *lctM*_543_AscI_RP	CTCAGCGGCGCGCCTGAAAGCTCATTTATCGC
pRSFDuet1-*lctM*-*procM* (pRSFDuet1-*lctM*-1-543-AscI)	Forward: *procM*_661_AscI_FP	CTTTCAGGCGCGCCAACGAGTCTGGAGGCAGC
	Reverse: *procM* _NotI_RP	AAGGAAAAAAGCGGCCGCTTATTCAGTAGGCCAGAGA

4.4.3 Construction of the expression plasmids for chimeric peptides

The generation of pET15b-*ProcALea_NisAStr* and pET15b-*ProcALea_LctAStr* constructs were described in section 2.4.2. The pET15b-*LctLea_ProcAStr* plasmids was constructed following the same procedures using primers and templates described in Table 4.3. The sequences of the inserts were checked by DNA sequencing. Unmodified CylLs peptide was obtained from Weixin Tang.

Table 4.3. Primer sequence used in section 4.4.3.

pET15b-*lctALea-procA3.3Str* (Template: pET15b-*lctA*, pET15b-*procA3.3*)	Forward-1: *lctA*_NdeI_FP	GGCAGCCATATGATGAAAGAACAAAACTCT
	Reverse-1: *lctA3.2Lea_ procA3.3Str* _Conn_RP	CTTATTTTAGGTGCAGGCGATACCGGCATC
	Forward-2: *lctA3.2Lea_ procA3.3Str* _Conn_FP	CTTATTTTAGGTGCAGGCGATACCGGCATC
	Reverse-2: *ProcA3.3*_XhoI_RP	ATACAATCCTCGAGCTATGCGCGGCACATTTTGG

4.4.4 Generation of dehydrated ProcA3.3 core peptide

Oxidized dehydrated ProcA3.3 peptide was generated using procedures described in section 3.4.3. An aliquot of 300 μL peptide was treated with 10 μL GluC (2 mg/mL) for 3 h, and the resulting peptide was injected to an analytical C18 HPLC column, and separated by HPLC using a gradient of 2% mobile phase B to 100% mobile phase B over 45 min (mobile phase A: water with 0.1% trifluoroacetic acid, mobile phase B: acetonitrile with 0.1% trifluoroacetic acid). The core peptide was collected and reduced with TCEP in the *in vitro* assays.

4.4.5 Construction of the expression plasmids for ProcA3.2 leader peptide

The plasmid pET15b-*procA3.2*[26] was isolated from *E. coli* DH5α cells and used as template for PCR reaction. The DNA sequence for ProcA3.2 leader was PCR amplified using Phusion polymerase by 30 cycles of denaturing (98 °C for 30 s), annealing (55 °C for 30 s), and extending (72 °C for 15 s) using forward primer and reverse primer indicated in Table 4.4. The PCR product was digested with NdeI and XhoI restriction enzymes and ligated pET15b vector to generate the construct pET15b-*procA3.2Lea*. The sequences of the inserts were checked by DNA sequencing.

Table 4.4. Primer sequence used in section 4.4.5.

pET15b-*procA3.2Lea* (Template: pET15b-*procA3.2)*	Forward: *procA3.2*_NdeI_FP	TCAGATCATATGATGTCAGA AGAACAACTC
	Reverse: *procA3.2Lea*_XhoI_RP	TCAGATCTCGAGCTATCCCC CAGCCACACCTTC

4.4.6 Labelling ProcA3.2 leader peptides with fluorescein isothiocyanate (FITC)

FITC (purchased from Sigma) was dissolved in DMF at 10 mg/mL. An aliquot of 1 mg His_6-tagged ProcA3.2 leader peptide was dissolved in 0.5 mL of conjugation buffer (50 mM borate buffer, pH 8.5). An excess amount of FITC was added (FITC concentration was 20-fold that of the peptide concentration) to the peptide solution and the mixture was incubated at room temperature in the dark for 24 h. The resulting peptide mixture was subjected to extensive dialysis in the 20 mM HEPES buffer (pH 7.5) to remove DMF and unreacted FITC, followed by repeated desalting using an Amicon Ultra-15 Centrifugal Filter Unit (3kDa MWCO) until all excess FITC was removed from the peptide solution.

4.4.7 Fluorescence polarization (FP) assay using FITC-ProcA3.2Lea

All FP experiments were performed on a Synergy H4 Hybrid plate reader (BioTek) using a 96-well black opaque half area plate (Costar 3694). Increasing amounts of LanM or truncated LanM enzymes were added to fluorescein-labeled peptides (10 or 20 nM) in LanM Start buffer (20 mM HEPES, 1 M NaCl, pH 7.5). Data analysis was performed using Origin 9.0, and curves were fitted to a hyperbolic equation to calculate K_D ($y = (a*x)/(K_D + x) + c$).

4.4.8 Experimental procedures listed in other chapters

Expression and purification of unmodified and modified peptides (section 2.4.5); Expression and purification of LanM enzymes (section 2.4.6);

In vitro enzymatic assay and endoprotease digestion (section 2.4.7);

Iodoacetamide assay (section 2.4.9);

Generation of dehydrated ProcA3.3 (section 3.4.3).

4.4.9 Calculated and observed masses in the mass spectra

Table 4.5. Calculated and observed masses in the mass spectra.

* monoisotopic masses plus one proton were shown for all peaks

Figure	Peak	Calculated mass	Observed mass
Figure 4.4 (A)	M - H_2O (Red)	2143.0	2142.5
	M (Red)	2161.0	2160.5
	M - 2 H_2O (Blue)	2125.0	2124.0
Figure 4.4 (B)	M - H_2O (Red)	2635.2	2634.6
	M (Red)	2653.2	2652.6
	M - 3 H_2O (Blue)	2599.2	2598.3
	M - 2 H_2O (Blue)	2617.2	2616.3
	M - H_2O (Blue)	2635.2	2634.6
Figure 4.4 (C)	M (Red)	7705.6	7703.3
	M (Blue)	7705.6	7703.6
Figure 4.5	M - 3 H_2O (Red)	2599.2	2599.6
	M - 3 H_2O + 1 IAA (Red)	2656.2	2656.6
	M - 3 H_2O (Blue)	2599.2	2599.0
	M - 3 H_2O + 1 IAA (Blue)	2656.2	2656.6
	M - 3 H_2O + 2 IAA (Blue)	2713.2	2713.1
Figure 4.7	M - 3 H_2O + 2 IAA (Red)	2713.2	2714.0
	M - 2 H_2O + 2 IAA (Red)	2731.2	2732.0
	M - 3 H_2O + 2 IAA (Blue)	2713.2	2713.2
	M - 2 H_2O + 2 IAA (Blue)	2731.2	2731.2
Figure 4.8 (A)	M - H_2O (Blue)	2143.0	2142.7
	M (Blue)	2161.0	2160.7
	M - 2 H_2O (Red)	2125.0	2124.9
Figure 4.8 (B)	M - H_2O (Blue)	2635.2	2635.0
	M (Blue)	2653.2	2654.0
	M - 3 H_2O (Blue)	2599.2	2599.9
	M - 2 H_2O (Blue)	2617.2	2617.9
	M - H_2O (Blue)	2635.2	2634.0

(Table 4.5 continued)

Figure	Peak	Calculated mass	Observed mass
Figure 4.11	M - H_2O (Red, Lct A core, 1 h)	2986.3	2986.5
	M (Red, Lct A core, 1 h)	3004.3	3004.6
	M - H_2O (Red, Lct A core, 3 h)	2986.3	2987.0
	M - 3 H_2O (Red, leader + core, 1 h)	2950.3	2950.8
	M - 3 H_2O (Red, leader + core, 3 h)	2950.3	2950.8
	M - H_2O (Blue, Lct A core, 1 h)	2986.3	2986.5
	M (Blue, Lct A core, 1 h)	3004.3	3004.5
	M - H_2O (Blue, Lct A core, 3 h)	2986.3	2986.7
	M - 3 H_2O (Blue, leader + core, 1 h)	2950.3	2950.7
	M - 3 H_2O (Blue, leader + core, 3 h)	2950.3	2950.7
Figure 4.12	M - 5 H_2O (Red)	8788.2	8791.9
	M - 4 H_2O (Red)	8806.2	8808.8
	M - 5 H_2O + PO_3^- (Red)	8868.2	8969.6
	M (Blue)	8878.2	8878.1
Figure 4.13	M - 3 H_2O (LctA)	2790.2	2790.1
	M (LctA)	2844.2	2844.1
	M - 3 H_2O + 1 IAA (LctA, IAA)	2847.2	2847.4
	M - 3 H_2O + 2 IAA (LctA, IAA)	2904.2	2904.4
	M + 3 IAA (LctA, IAA)	3015.2	3015.5
	M (ProcALea-LctAstr)	2844.2	2844.2
	M (ProcA3.3)	2653.2	2653.1
Figure 4.14	M - H_2O (ProcA3.3)	2635.2	2635.1
	M (ProcA3.3)	2653.2	2653.0
	M - H_2O + 1 IAA (ProcA3.3, IAA)	2692.2	2691.5
	M + 2 IAA (ProcA3.3, IAA)	2767.2	2767.5
	M (ProcALea-LctAstr)	2844.2	2843.9
	M (LctA)	2844.2	2843.8

4.5 REFERENCES

1. Knerr, P. J., and van der Donk, W. A. (2012) Discovery, biosynthesis, and engineering of lantipeptides, *Annu. Rev. Biochem. 81*, 479-505.

2. Arnison, P. G., Bibb, M. J., Bierbaum, G., Bowers, A. A., Bugni, T. S., Bulaj, G., Camarero, J. A., Campopiano, D. J., Challis, G. L., Clardy, J., Cotter, P. D., Craik, D. J., Dawson, M., Dittmann, E., Donadio, S., Dorrestein, P. C., Entian, K. D., Fischbach, M. A., Garavelli, J. S., Goransson, U., Gruber, C. W., Haft, D. H., Hemscheidt, T. K., Hertweck, C., Hill, C., Horswill, A. R., Jaspars, M., Kelly, W. L., Klinman, J. P., Kuipers, O. P., Link, A. J., Liu, W., Marahiel, M. A., Mitchell, D. A., Moll, G. N., Moore, B. S., Muller, R., Nair, S. K., Nes, I. F., Norris, G. E., Olivera, B. M., Onaka, H., Patchett, M. L., Piel, J., Reaney, M. J., Rebuffat, S., Ross, R. P., Sahl, H. G., Schmidt, E. W., Selsted, M. E., Severinov, K., Shen, B., Sivonen, K., Smith, L., Stein, T., Sussmuth, R. D., Tagg, J. R., Tang, G. L., Truman, A. W., Vederas, J. C., Walsh, C. T., Walton, J. D., Wenzel, S. C., Willey, J. M., and van der Donk, W. A. (2013) Ribosomally synthesized and post-translationally modified peptide natural products: overview and recommendations for a universal nomenclature, *Nat. Prod. Rep. 30*, 108-160.

3. Chatterjee, C., Paul, M., Xie, L., and van der Donk, W. A. (2005) Biosynthesis and mode of action of lantibiotics, *Chem. Rev. 105*, 633-684.

4. Li, B., and van der Donk, W. A. (2007) Identification of essential catalytic residues of the cyclase NisC involved in the biosynthesis of nisin, *J. Biol. Chem. 282*, 21169-21175.

5. Xie, L., Miller, L. M., Chatterjee, C., Averin, O., Kelleher, N. L., and van der Donk, W. A. (2004) Lacticin 481: in vitro reconstitution of lantibiotic synthetase activity, *Science 303*, 679-681.

6. Piard, J. C., Muriana, P. M., Desmazeaud, M. J., and Klaenhammer, T. R. (1992) Purification and partial characterization of lacticin 481, a lanthionine-containing bacteriocin produced by Lactococcus lactis subsp. lactis CNRZ 481, *Appl. Environ. Microbiol. 58*, 279-284.

7. Paul, M., Patton, G. C., and van der Donk, W. A. (2007) Mutants of the zinc ligands of lacticin 481 synthetase retain dehydration activity but have impaired cyclization activity, *Biochemistry 46*, 6268-6276.

8. You, Y. O., and van der Donk, W. A. (2007) Mechanistic investigations of the dehydration reaction of lacticin 481 synthetase using site-directed mutagenesis, *Biochemistry 46*, 5991-6000.

9. Levengood, M. R., Knerr, P. J., Oman, T. J., and van der Donk, W. A. (2009) In vitro mutasynthesis of lantibiotic analogues containing nonproteinogenic amino acids, *J. Am. Chem. Soc. 131*, 12024-12025.

10. Oman, T. J., Knerr, P. J., Bindman, N. A., Velasquez, J. E., and van der Donk, W. A. (2012) An engineered lantibiotic synthetase that does not require a leader peptide on its substrate, *J. Am. Chem. Soc. 134*, 6952–6955.

11. Knerr, P. J., Oman, T. J., Garcia de Gonzalo, C., Lupoli, T. J., S., W., and van der Donk, W. A. (2012) Non-Proteinogenic Amino Acids in Lacticin 481 Analogues Result in More Potent Inhibition of Peptidoglycan Transglycosylation, *ACS Chem. Biol. 7*, 1791-1795.

12. Oman, T. J., and van der Donk, W. A. (2009) Insights into the Mode of Action of the Two-Peptide Lantibiotic Haloduracin, *ACS Chem. Biol. 4*, 865-874.

13. Dufour, A., Hindré, T., Haras, D., and Le Pennec, J. P. (2007) The biology of the lantibiotics of the lacticin 481 subgroup is coming of age, *FEMS Microbiol. Rev. 31*, 134-167.

14. Hsu, S. T., Breukink, E., Bierbaum, G., Sahl, H. G., de Kruijff, B., Kaptein, R., van Nuland, N. A., and Bonvin, A. M. (2003) NMR study of mersacidin and lipid II interaction in dodecylphosphocholine micelles. Conformational changes are a key to antimicrobial activity, *J. Biol. Chem. 278*, 13110-13117.

15. Bottiger, T., Schneider, T., Martinez, B., Sahl, H. G., and Wiedemann, I. (2009) Influence of Ca(2+) ions on the activity of lantibiotics containing a mersacidin-like lipid II binding motif, *Appl. Environ. Microbiol. 75*, 4427-4434.

16. Islam, M. R., Nishie, M., Nagao, J., Zendo, T., Keller, S., Nakayama, J., Kohda, D., Sahl, H. G., and Sonomoto, K. (2012) Ring A of nukacin ISK-1: a lipid II-binding motif for type-A(II) lantibiotic, *J. Am. Chem. Soc. 134*, 3687-3690.

17. Gilmore, M. S., Segarra, R. A., Booth, M. C., Bogie, C. P., Hall, L. R., and Clewell, D. B. (1994) Genetic structure of the *Enterococcus faecalis* plasmid pAD1-encoded cytolytic toxin system and its relationship to lantibiotic determinants, *J. Bacteriol. 176*, 7335-7344.

18. Cox, C. R., Coburn, P. S., and Gilmore, M. S. (2005) Enterococcal cytolysin: a novel two component peptide system that serves as a bacterial defense against eukaryotic and prokaryotic cells, *Curr. Protein Pept. Sci. 6*, 77-84.

19. Lawton, E. M., Ross, R. P., Hill, C., and Cotter, P. D. (2007) Two-peptide lantibiotics: a medical perspective, *Mini Rev. Med. Chem. 7*, 1236-1247.

20. Huycke, M. M., Spiegel, C. A., and Gilmore, M. S. (1991) Bacteremia caused by hemolytic, high-level gentamicin-resistant *Enterococcus faecalis*, *Antimicrob. Agents Chemother. 35*, 1626-1634.

21. Ike, Y., and Clewell, D. B. (1992) Evidence that the hemolysin/bacteriocin phenotype of Enterococcus faecalis subsp. zymogenes can be determined by plasmids in different

incompatibility groups as well as by the chromosome, *J. Bacteriol. 174*, 8172-8177.

22. Coburn, P. S., and Gilmore, M. S. (2003) The Enterococcus faecalis cytolysin: a novel toxin active against eukaryotic and prokaryotic cells, *Cell Microbiol. 5*, 661-669.

23. Gilmore, M. S., Skaugen, M., and Nes, I. (1996) Enterococcus faecalis cytolysin and lactocin S of Lactobacillus sake, *Antonie van Leeuwenhoek 69*, 129-138.

24. Tang, W., and van der Donk, W. A. (2013) The sequence of the enterococcal cytolysin imparts unusual lanthionine stereochemistry, *Nat. Chem. Biol. 9*, 157-159.

25. Tang, W., and van der Donk, W. A. (2012) Structural characterization of four prochlorosins: a novel class of lantipeptides produced by planktonic marine cyanobacteria, *Biochemistry 51*, 4271-4279.

26. Li, B., Sher, D., Kelly, L., Shi, Y., Huang, K., Knerr, P. J., Joewono, I., Rusch, D., Chisholm, S. W., and van der Donk, W. A. (2010) Catalytic promiscuity in the biosynthesis of cyclic peptide secondary metabolites in planktonic marine cyanobacteria, *Proc. Natl. Acad. Sci. U.S.A. 107*, 10430-10435.

27. Zhang, Q., Yu, Y., Velasquez, J. E., and van der Donk, W. A. (2012) Evolution of lanthipeptide synthetases, *Proc. Natl. Acad. Sci. U.S.A. 109*, 18361-18366.

28. Zhang, Q., Yang, X., Wang, H., and van der Donk, W. A. (2014) High divergence of the precursor peptides in combinatorial lanthipeptide biosynthesis, *ACS Chem. Biol. 9*, 2686-2694.

29. Li, B., Yu, J. P., Brunzelle, J. S., Moll, G. N., van der Donk, W. A., and Nair, S. K. (2006) Structure and mechanism of the lantibiotic cyclase involved in nisin biosynthesis, *Science 311*, 1464-1467.

30. Coordinators, N. R. (2015) Database resources of the National Center for Biotechnology Information, *Nucleic Acids Res. 43*, D6-17.

31. Lubelski, J., Khusainov, R., and Kuipers, O. P. (2009) Directionality and Coordination of Dehydration and Ring Formation during Biosynthesis of the Lantibiotic Nisin, *J. Biol. Chem. 284*, 25962-25972.

32. Lee, M. V., Ihnken, L. A., You, Y. O., McClerren, A. L., van der Donk, W. A., and Kelleher, N. L. (2009) Distributive and directional behavior of lantibiotic synthetases revealed by high-resolution tandem mass spectrometry, *J. Am. Chem. Soc. 131*, 12258-12264.

33. Thibodeaux, C. J., Ha, T., and van der Donk, W. A. (2014) A Price To Pay for Relaxed Substrate Specificity: A Comparative Kinetic Analysis of the Class II Lanthipeptide Synthetases ProcM and HalM2, *J. Am. Chem. Soc. 136*, 17513-17529.

34. Jungmann, N. A., Krawczyk, B., Tietzmann, M., Ensle, P., and Süssmuth, R. D. (2014) Dissecting Reactions of Nonlinear Precursor Peptide Processing of the Class III Lanthipeptide Curvopeptin, *J. Am. Chem. Soc. 136*, 15222-15228.

35. Mukherjee, S., and van der Donk, W. A. (2014) Mechanistic Studies on the Substrate-Tolerant Lanthipeptide Synthetase ProcM, *J. Am. Chem. Soc. 136*, 10450-10459.

36. Yu, Y., Mukherjee, S., and van der Donk, W. A. (2015) Product Formation by the Promiscuous Lanthipeptide Synthetase ProcM is under Kinetic Control, *J. Am. Chem. Soc. 137*, 5140-5148.

37. Goto, Y., Li, B., Claesen, J., Shi, Y., Bibb, M. J., and van der Donk, W. A. (2010) Discovery of unique lanthionine synthetases reveals new mechanistic and evolutionary insights, *PLoS Biol. 8*, e1000339.

38. Shimafuji, C., Noguchi, M., Nishie, M., Nagao, J. I., Shioya, K., Zendo, T., Nakayama, J., and Sonomoto, K. (2015) In vitro catalytic activity of N-terminal and C-terminal domains in NukM, the post-translational modification enzyme of nukacin ISK-1, *J. Biosci. Bioeng.*, doi: 10.1016.

CHAPTER V: YEAST SURFACE DISPLAY OF LANTHIPEPTIDE LEADER PEPTIDES

5.1 INTRODUCTION

As described in chapter I, the precursor peptides of lanthipeptides are ribosomally translated and are composed of an N-terminal leader peptide and a C-terminal core peptide. During the biosynthesis of lanthipeptides, post-translational modifications occur in the core peptides, followed by the removal of the leader peptides to generate the mature product. The leader peptides are believed to carry out the following functions: First, keep the modified core peptide inactive in order to protect the producing strains;[1-3] Second, act as a secretion signal to be recognized by the lanthipeptide transporters;[2, 3] Third, serve as a recognition motif for the biosynthetic enzymes and guide the process of the post-translational modifications;[3-6] And fourth, bind to the biosynthetic enzymes and shift the equilibrium between the inactive and active enzyme towards the latter form.[3, 7]

The leader peptides of different lanthipeptides preserve a certain degree of sequence conservation, especially the protease recognition motif just before the core region.[3] The leader peptides of class I lanthipeptides are about 25 amino acids in length and are rich in aspartate residues, and contain a conserved FNLD box.[8] The leader peptides of class II lanthipeptides are also rich in aspartate and glutamate residues, contain an ELXXBX motif (B = V, L or I), and usually end in a double glycine motif. A recent paper reported the first co-crystal structure of a lanthionine synthetase (NisB) in complex with its substrate peptide (NisA),[9] which provides insight into the substrate recognition of this group of enzymes. In the co-crystal

structure, the FNLD motif of NisA interacts with an anti-parallel β strand motif of NisB, and this scheme of substrate binding might be general among the lanthionine synthetases.

There are several motivations to engineer lanthipeptide leader peptides for tighter binding affinity to the modification enzymes. The first reason is to facilitate crystallographic studies of the lanthipeptide synthetases. When this project was carried out, there was no crystal structure reported for the LanB dehydratases and the bifunctional LanM synthetases although extensive efforts were made. The cyclase NisC in the nisin biosynthetic pathway is one of the few lanthipeptide modification enzymes for which a crystal structure has been solved. The crystal structure of NisC was solved to a resolution of 2.5 Å and efforts were made to obtain a co-crystal structure of NisC with its substrate in its binding pocket.[10] The difficulties in obtaining co-crystal structures of the lanthionine synthetase and the substrate might be due to the relatively weak binding affinity. In order to solve this problem, yeast surface display was selected to evolve the lanthipeptide leader peptides for tighter binding, which might also be useful to obtain crystals of the lanthionine synthetases without reported structures.

A previous study in our lab carried out by Dr. Oman showed that LctM fused with the LctA leader peptide (ConFusion LctM) catalyzed the modification of the LctA core peptide without its leader peptide attached to it.[11] However, the catalytic efficiency was dramatically decreased compared to the catalytic efficiency of wild type LctM on its natural substrate. It was envisioned that if LctM is fused with a more tightly bound LctA leader peptide, the leader peptide might be able to trap the enzyme in its active form and increase the catalytic efficiency of the fusion enzyme. We are interested in engineering ConFusion LctM because by utilizing this enzyme, lacticin 481 could be produced *in vitro* efficiently and the purification process

would be largely simplified. This new strategy could also be applied to the production of other class II lanthipeptides, and to synthetic core peptides containing non-natural amino acids.

As described in chapter II, ProcM was able to dehydrate and cyclize non-natural peptides attached to the ProcA leader peptide.[12] This unusual catalytic promiscuity of ProcM was utilized for cyclic peptide library generation (work of lab member Xiao Yang). However, not all chimeric peptides were modified completely by ProcM. The modification efficiency, especially the cyclization efficiency was not satisfactory for some of the substrates. Considering leader peptides may act as an activator for the modification enzymes, engineering the ProcA leader peptide may be helpful in terms of solving this problem. It is possible that if the binding affinity of the leader peptide to the enzyme is too high, the substrate will not be released by the enzyme. But it may be feasible to avoid this problem by engineering a leader peptide with a desired binding affinity that meets both requirements.

5.2 RESULTS

5.2.1 The yeast surface display system

In the yeast surface display system, lanthipeptide libraries are displayed on the surface of yeast cells by fusing to the yeast surface protein AGA-2 (Figure 5.1).[13] The fusion proteins are recruited to the yeast surface by the membrane bound AGA-1 protein and disulfide bonds are formed between these two partners. To detect the binding of lanthipeptides and their modifying enzymes, the yeast libraries were incubated with purified enzymes containing either an epitope tag or a EGFP (enhanced green fluorescent protein)-tag. FACS (fluorescence-activated cell sorting) analysis of the yeast libraries using flow cytometry was used to identify and separate yeast cells bound to

fluorescently labelled enzyme. To engineer lanthipeptide leader peptides with stronger binding affinity, yeast libraries of lanthipeptides with diversified leader regions were generated. Because the detection of the initial binding between the truncated leader peptides and the modification enzymes were not achieved easily, the full length lanthipeptides were displayed to increase the initial binding to the modification enzymes and facilitate the sorting process. For the ProcA library, only the leader peptides were displayed because the ProcA leader peptides were relatively long and there was no sequence conservation between the core peptides.[14]

To construct the desired lanthipeptide peptide libraries, the first step is to insert the genes encoding lanthipeptides after the AGA-2 gene in the yeast display vector pCT302 (Figure 5.1.B and section 5.4.2). The presence of HA and c-myc tags at both ends of the lanthipeptide gene enabled the detection of the lanthipeptide expression by immunostaining. Error prone PCR using a long reverse primer was carried out using these constructs to introduce random mutations into the leader region only (Figure 5.1.C).

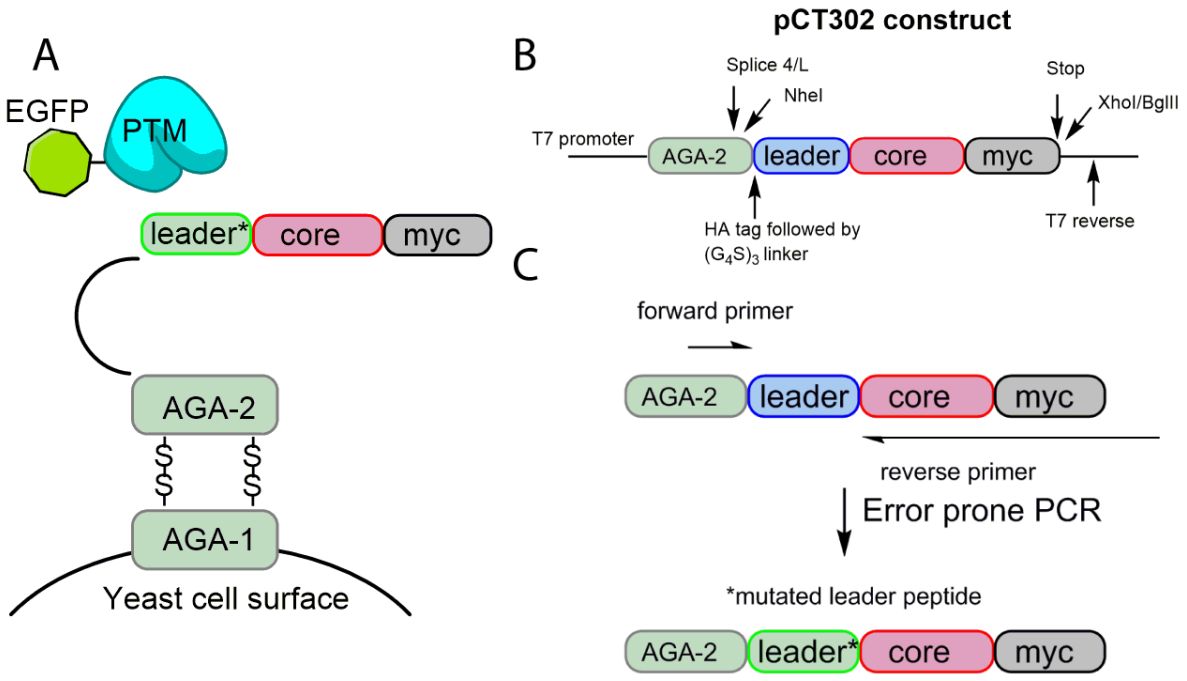

Figure 5.1. The yeast surface system: (A) Diagram of a lanthipeptide library with diversified leader regions displayed on the surface of yeast cells fused to Aga-2. A detection scheme involving EGFP-tagged post-translational modification enzyme (PTM) bound to the lanthipeptide library may be used to evaluate the binding affinity; (B) the pCT302 construct of the AGA-2-lanthipeptide fusion protein showing the arrangement of the gene elements; (C) the primer design for the generation of the lanthipeptide library with a diversified leader region.

5.2.2 The evaluation and sorting of lanthipeptide yeast libraries

The yeast libraries of three lanthipeptides (NisA, LctA and ProcA3.2Lea) with diversified leader regions were generated as discussed in section 5.2.1. To evaluate the error rate of the libraries, the error prone PCR products were ligated into a T-vector and used to transform *E. coli* DH5α cells. For each library, 20 colonies were picked for plasmid extraction and DNA sequencing. The sequencing results were used to estimate the error rate in the leader region.

For all the lanthipeptide libraries, an error rate of 0.5% was achieved using protocols descried in section 5.4.3. The yeast cell library construction and sorting processes are described in detail in section 5.4.4-5.4.8.

For the NisA library, all the isolated yeast colonies from the sorting process contained peptide genes with shifts in the open reading frame, which resulted in the expression of unrelated peptides. For the ProcA3.2Lea library, only colonies with false positive signals were sorted. Thus, further selection was not carried out for these two libraries. A subset of 60 colonies (yeast colony-LctA-1 to 60) were selected from the 6^{th} round of sorting of the first generation of LctA library. After the induction of peptide overexpression, yeast cell cultures containing individual plasmids (pCT302-LctA-1 to 60) were stained with EGFP-LctM and the mean fluorescence intensity (MFI) was measured for each yeast colony. Plasmids were isolated from colonies with increased MFI compared to a colony with WT pCT302-LctA and were used to transform *E. coli* DH10B cells. Plasmids isolated from DH10B cells were used for DNA sequencing. Some of these plasmids had identical LctA sequences, and plasmids with different DNA sequences were used to transform fresh electrocompetent yeast cells. The reason for transforming fresh yeast cells is to avoid any mutations in the yeast genome that might contribute to the enhanced fluorescence signal in the sorting process. Yeast colony-LctA-34 was isolated from the above procedures, which had an N8D mutation in the leader region (Figure 5.2). The pCT302-LctA-34 plasmid was used as the template for the construction of the second generation LctA library. The sorting and characterization processes were the same as for the first generation library, and four LctA mutants (LctA-34-7, LctA-34-11, LctA-34-13, and LctA-34-28) were isolated, which all had the Q4L, N5D, N8D, and S16D mutations

(Figure 5.2). Yeast colony-LctA-34-28 showed the highest MFI and the pCT302-LctA-34-28 plasmid was used as template for the construction of the third generation LctA library. However, the isolated plasmids from this library contained no mutation in the leader region, instead, they all contained mutations resulting in open reading frame shift. As a result, no further sorting processes were carried out.

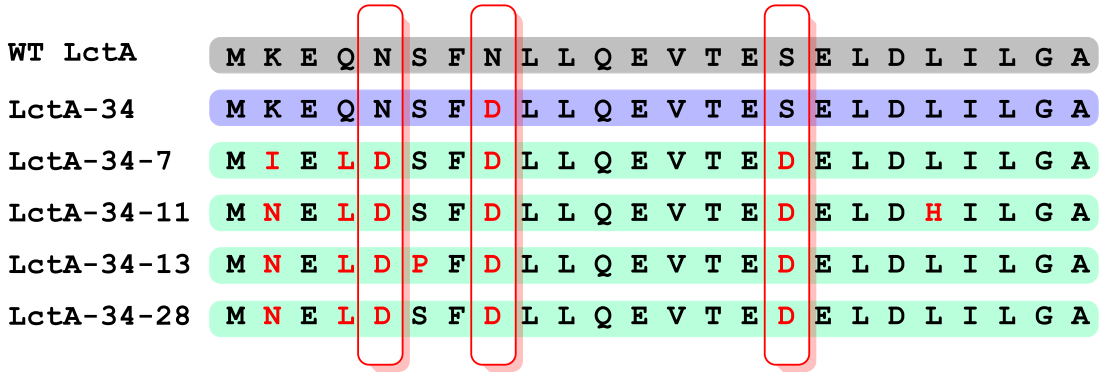

Figure 5.2. The amino acid sequences of WT LctA and engineered LctA peptides obtained from the sorting of the LctA library. The mutations are labelled in red, and the three conserved aspartic acid mutations are highlighted with red box.

5.2.3. Yeast colonies with the engineered LctA lanthipeptide bind stronger to EGFP-LctM

As described in section 5.2.2, the isolated plasmids obtained from the yeast library sorting were used to transform fresh electrocompetent yeast cells, and the yeast surface expression of the engineered LctA peptides was induced. The binding affinity of LctA analogs to LctM was first evaluated by FACS analysis to measure the MFI of the yeast population after staining with EGFP-LctM. The MFI values of the best binders isolated from the first and second generation of the LctA library were compared to a yeast colony displaying WT LctA. Yeast colony-LctA-

34 showed about 10-fold increase in MFI compared to the yeast colony with WT peptide, while yeast colony-LctA-34-28 showed a 30-fold increase (Figure 5.3). This result ruled out the possibility that mutations in the yeast genome resulted in false positive selection, and suggested that the binding affinity of engineered LctA peptide (LctA-34-28) to LctM was improved compared to the WT peptide.

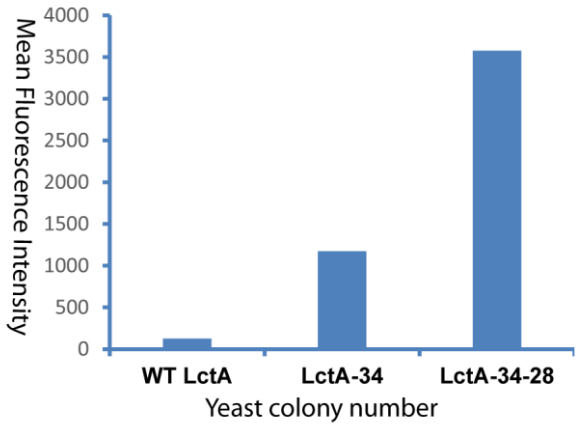

Figure 5.3. The mean fluorescence intensity for yeast colony displaying WT LctA and engineered LctA peptides after staining with 3.6 μM EGFP-LctM.

5.2.4 The engineered LctA leader peptide binds stronger to LctM than WT leader peptide

The MFI of the yeast colony could be influenced by other factors besides the binding affinity of the expressed peptide to the fluorescently labelled protein, such as expression level of the peptides and non-specific binding to the EGFP tag. In order to directly measure the binding affinity between the peptide and the modification enzymes, several *in vitro* experiments using purified peptide and enzymes were carried out. Isothermal titration calorimetry (ITC) experiments were first attempted, however, the LctM enzyme precipitated

in the metal sample cell in all experiments. Therefore, another technique, fluorescence polarization (FP), was applied. First, the WT LctA leader peptide (obtained from Dr. Champak Chatterjee, generated using solid-phase peptide synthesis) was fluorescently labeled with fluorescein isothiocyanate (FITC) which primarily reacts with amine groups (described in section 5.4.11 and 5.4.12). The ESI-MS (Figure 5.4A) and analytical HPLC (Figure 5.4B) showed the purity of the product FITC-LctALea, and the tandem MS data showed that the FITC group was attached to the N-terminal amine group specifically and not to the amine group of the lysine residue at the second position (Figure 5.4C). LctA-34-28 leader peptide (ordered from United Peptide) was also labelled with FITC and purified using the same protocol.

Fluorescence polarization experiments (section 5.4.13) using FITC-LctALea peptide and purified LctM showed a K_D value of around 5.2 μM (Figure 5.5). As the activity of LctM requires ATP,[15] the binding affinity of FITC-LctALea was also measured in the presence of the non-hydrolysable ATP analog (AMP-PNP) at increasing concentration (Figure 5.5). But no significant changes in the K_D or any correlations of the K_D to AMP-PNP concentration were observed, which indicates that the binding of LctA leader peptide to LctM is not likely to be ATP-dependent. The K_D of FITC-LctALea and FITC-LctALea-34-28 to LctM were measured by FP experiments (section 5.4.13), and the engineered LctA leader peptide showed a 4-5 fold stronger binding affinity (Figure 5.6).

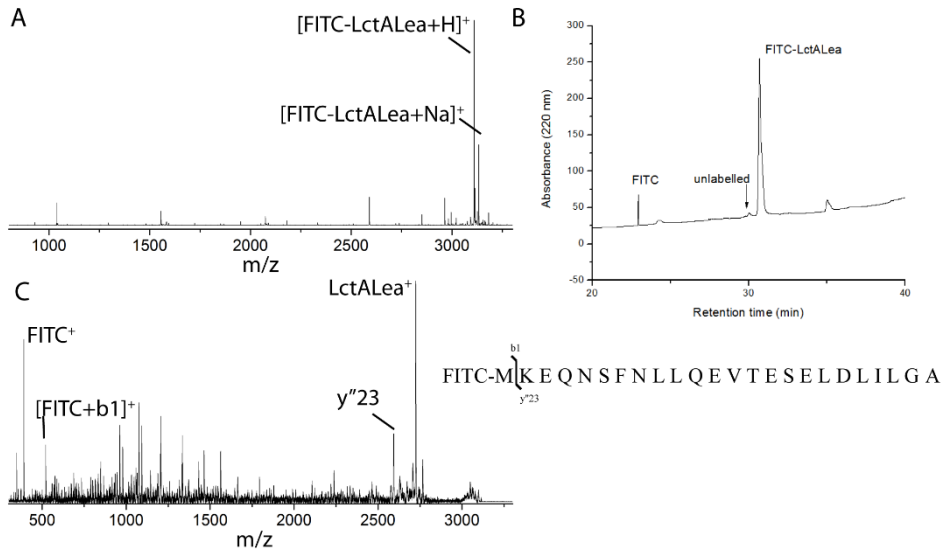

Figure 5.4. Analysis of the FITC-LctALea peptide: (A) ESI-MS showing the observed mass of FITC-LctALea (calculated mass: 3109.8, observed mass: 3110.0); (B) analytical HPLC showing the unlabeled peptide was removed in the purification process; (C) tandem MS data showing that the FITC group was attached to the N-terminal free amine group.

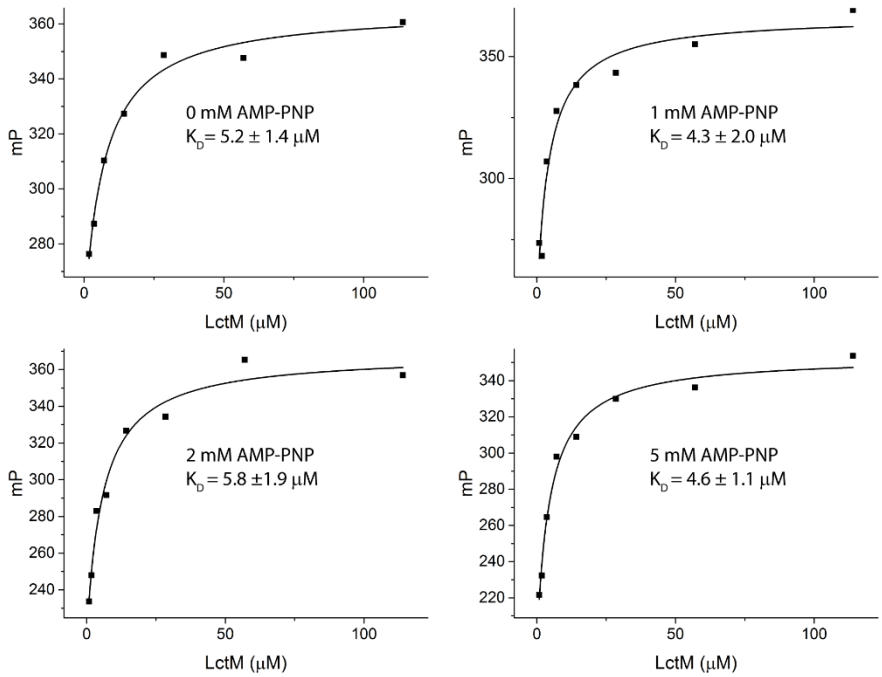

Figure 5.5. Fluorescence polarization assays with LctM using 20 nM FITC-LctALea in the presence of 0 mM, 1 mM，2 mM and 5 mM AMP-PNP.

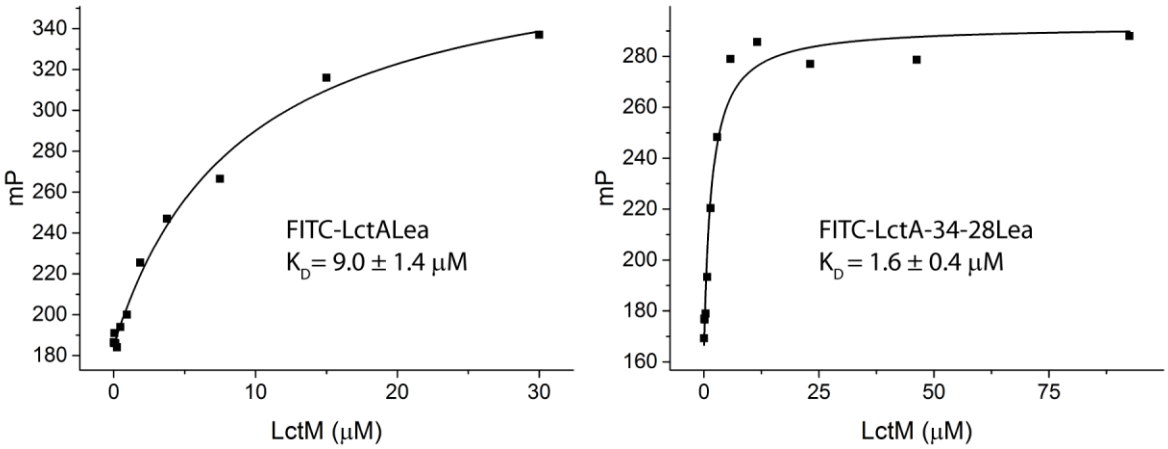

Figure 5.6. Fluorescence polarization assays with LctM using 5 nM FITC-LctALea or FITC-LctALea-34-28.

To minimize any interference of the bulky FITC group, WT LctA leader peptide and engineered LctA leader peptide were also labelled with FITC-Ahx at the N-terminus (Ahx is a linear six carbon aminohexanoic group inserted between the FITC group and the leader peptide). FITC-Ahx-LctALea and FITC-Ahx-LctALea-34-28 were purchased from United Peptide and the purity was > 95% for both peptides. The K_D of FITC-Ahx-LctALea and FITC-Ahx-LctALea-34-28 binding to LctM were measured using the same protocol, and again a 4~5 fold increase in the binding affinity was observed for the engineered LctA leader peptide (Figure 5.7). [16]

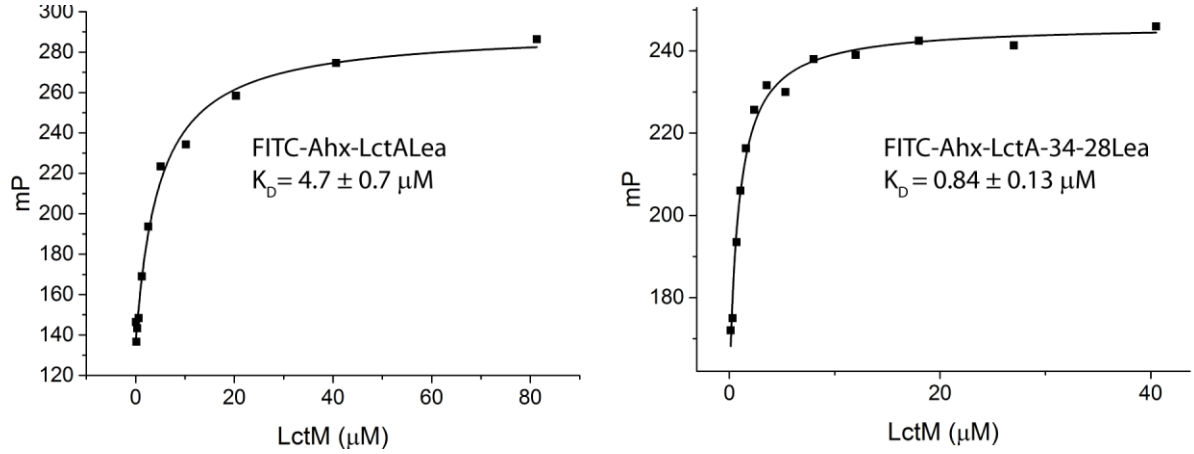

Figure 5.7. Fluorescence polarization assays with LctM using 5 nM FITC-Ahx-LctALea or FITC-Ahx-LctALea-34-28.

However, LctA peptides were shown to form soluble aggregation[16] and the peptide aggregates are likely to bind to LctM with different binding affinity and interfere with the binding assay. To avoid this problem, MBP-tagged LctALea and LctALea-34-28 peptides were constructed (section 5.4.10) and used in competitive FP assays. These proteins were analyzed by FPLC using an analytical size exclusion column, and it was shown that they existed as a stable protein monomer. MBP-LctALea-34-28 showed ~3 fold smaller K_D value (Figure 5.8), which is consistent with the results obtained by direct titration with FITC-labeled peptides.

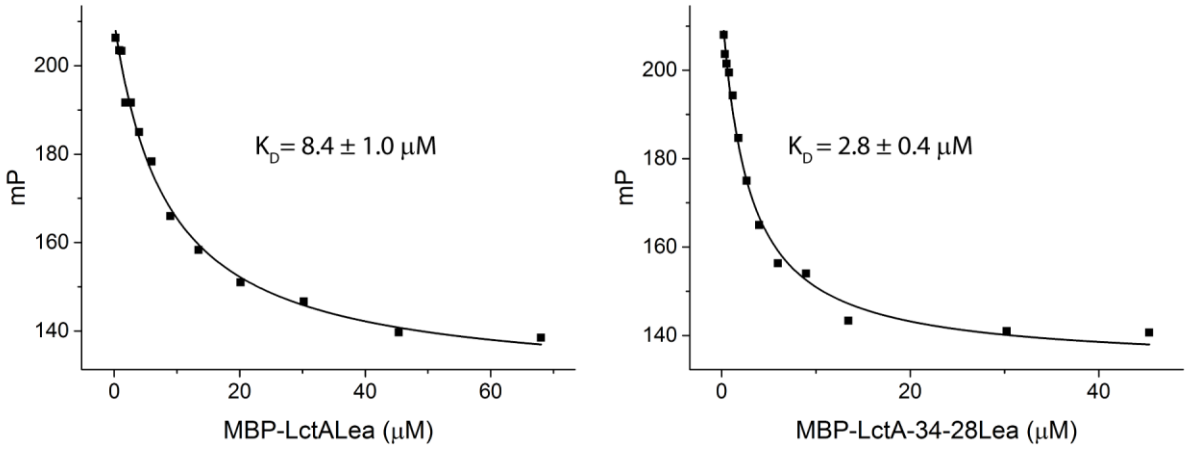

Figure 5.8. Competitive fluorescence polarization binding assays with 1.7 μM LctM using MBP-LctALea or MBP-LctA-34-28Lea to displace FITC-Ahx-LctA-34-28Lea (20 nM).

5.2.5 LctA containing the engineered leader peptide binds stronger to LctM

To measure the binding affinity of full length LctA and LctA-34-28, competitive fluorescence polarization assays (section 5.4.13) were carried out. His_6-LctA and His_6-LctA-34-28 (constructed as described in section 5.4.10) were used to displace FITC-Ahx-34-28 in the LctM binding assay, and the engineered full length LctA also showed 4-5 fold increase in the binding affinity (Figure 5.9). Similarly to section 5.2.4, the same displacement assay was carried out using MBP-tagged peptides. However, MBP-tagged full length LctA and LctA-34-28 existed as aggregates, and the FP data was not successfully fitted to a hyperbolic curve to calculate the binding affinity.

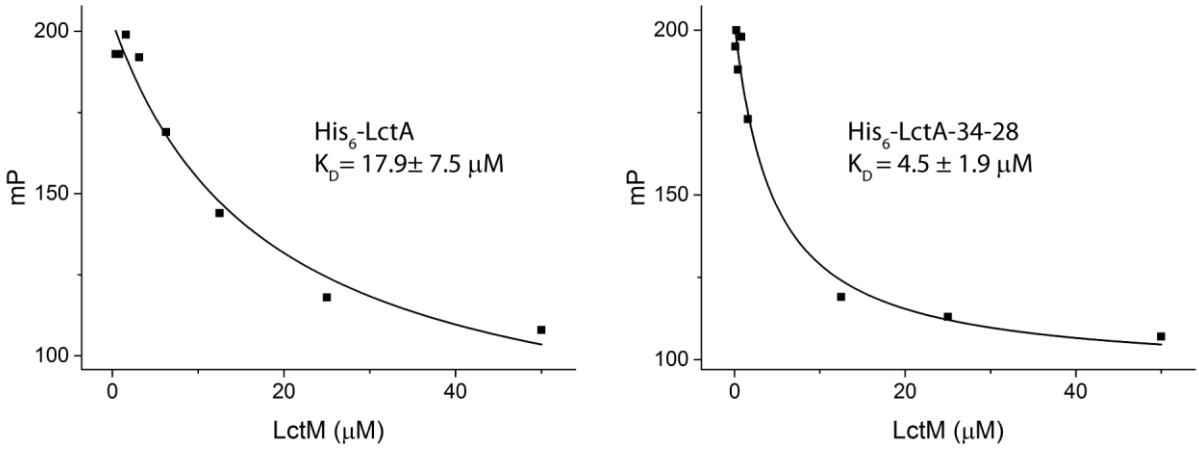

Figure 5.9. Competitive fluorescence polarization binding assays with LctM using His_6-LctA (red) or His_6-LctA-34-28 (blue) to displace FITC-Ahx-LctA-34-28Lea (20 nM).

5.2.6 The engineered LctA leader peptide binds stronger to the LctM dehydratase domain

As described in chapter IV, the two domains of LanM enzymes are active independently and the activities of both the dehydratase and cyclase domains are leader-dependent, which suggests two possible leader binding sites in LanMs.[17] The binding affinity of FITC-Ahx-LctALea and FITC-Ahx-LctALea-34-28 to LctM-1-543 (putative dehydratase domain of LctM) were measured. Similar to the observations with WT LctM, the engineered LctA leader peptide bound 4-5 fold stronger (Figure 5.10). It is of interest to see whether the two putative leader peptide binding sites prefer the same leader peptide sequences and whether both have increased affinity for the engineered leader peptide. However, this experiment could not be carried out because of the insolubility of LctM-544-922 (putative cyclase domain). MBP tagged LctM-544-922 was envisioned to increase the solubility, however, no overexpression of MBP-LctM-544-922 was observed.

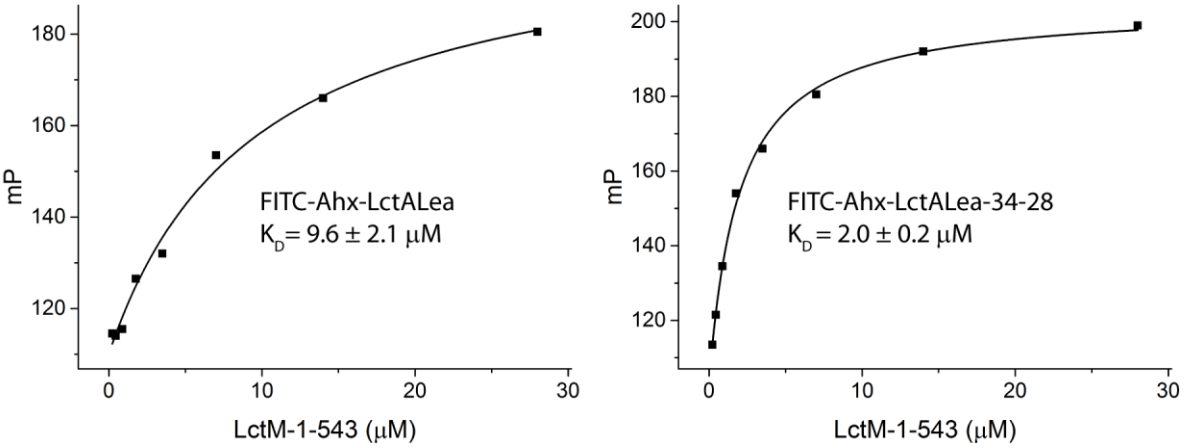

Figure 5.10. Fluorescence polarization assays with LctM-1-543 using 5 nM FITC-Ahx-LctALea or FITC-Ahx-LctALea-34-28.

5.2.7 The engineered LctA is a more efficient substrate for LctM

There are several potential applications of the tighter binding LctA leader peptide: 1. the possible generation of a co-crystal structure of LctM in complex with its substrate; 2, an improved ConFusion LctM (section 5.1) with higher catalytic efficiency; 3, a more efficient substrate for LctM. The first application has not been achieved so far due to the lack of crystal structure of either the full length LctM or the dehydratase domain of LctM. To see if engineered leader peptides could be applied for the second purpose, ConFusion LctM-34 and ConFusion LctM-34-28 were generated as described in section 5.4.9 and their activities were tested in assays with the LctA core peptide (section 5.4.14). The 4-fold dehydrated core peptide and starting material were quantified by the ion counts of the corresponding peak area in ESI-LC-MS experiments and normalized to be compared with the WT ConFusion-LctM. As shown in Figure 5.11, although ConFusion LctM-34 showed about 20% increase in the product conversion, ConFusion-34-28 showed about 56% decrease in efficiency. This result is not understood at present.

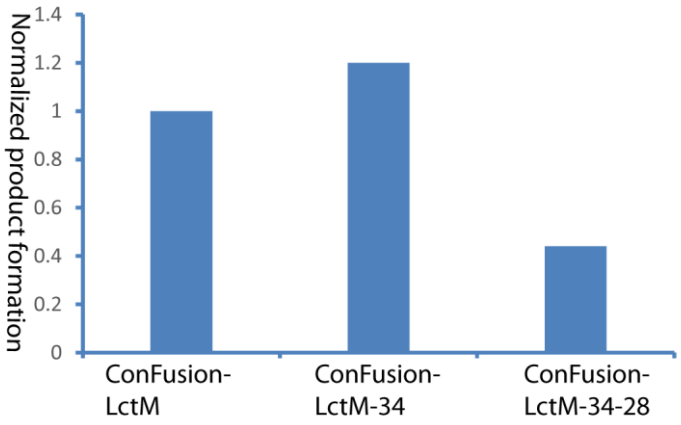

Figure 5.11. The relative product formation efficiency of ConFusion LctM enzymes in the assays with the LctA core peptide.

Next, the catalytic efficiency of LctM was studied using WT LctA and LctA-34-28. *In vitro* enzymatic assays were carried out using 25 μM peptides and 1 μM LctM, the reactions were quenched at different time points, treated with LysC to remove the leader peptide and spotted on a bioactivity assay plate to detect the formation of final product. As shown in Figure 5.12, the modification of LctA-34-28 was nearly completed within 5 min in the conditions used, but for the WT LctA, less product was formed even at 45 min.

A significant improvement in the efficiency of product turnover was observed with the engineered LctA peptide which is promising in the application of lanthipeptide production, however, it is difficult to tell whether this improvement is caused by increased catalytical efficiency or less aggregation of the substrate with three additional negatively charged residues, or a combination of both.

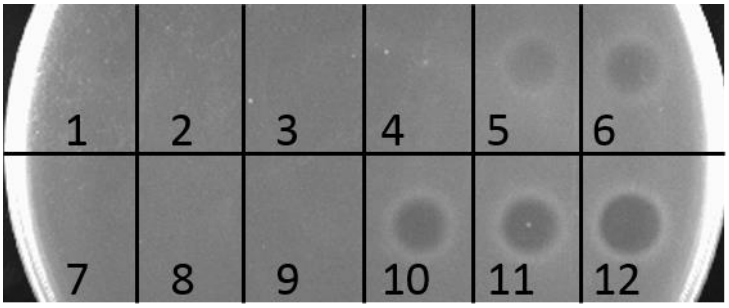

Figure 5.12. Antimicrobial assays against *L. lactis* HP. An aliquot of 10 μL of the following samples were spotted: (1-6) 25 μM WT LctA incubated with 1 μM LctM for 15 s, 1, 2, 5, 10 and 45 min, followed by the LysC digestion to remove the leader; (7-12) LctA-34-28 treated under the same conditions.

We have experienced challenges in analyzing the product conversion by mass spectrometry due to the poor ionization efficiency of the unmodified WT LctA substrate. The product to substrate ratio was often largely over-estimated for the WT LctA, but for the engineered LctA, this problem was much less severe. The bias in product/substrate ionization makes it impossible to compare the product formation of these two substrate by MS techniques accurately. But still, the *in vitro* enzymatic assays using LctA-34-28 showed a moderate increase in product formation in the mass spectra than assays using WT LctA (Figure 5.13). Efforts were made to eliminate the aggregation of the substrate and compare the actual rate of production formation without the interference of the aggregates. However, LctA and LctA-34-28 peptides with MBP, EGFP or SUMO tag at the N-terminal still existed as soluble protein aggregates. Since MBP-tagged LctA leaders exist as stable monomers, it is possible that peptide monomers could be obtained from the MBP-tagged LctA peptides with shorter core region. Therefore, MBP-LctA-1-38 and MBP-LctA-34-28-1-38 were constructed as described in section 5.4.10, however, these two proteins were not successfully overexpressed in *E. coli.*

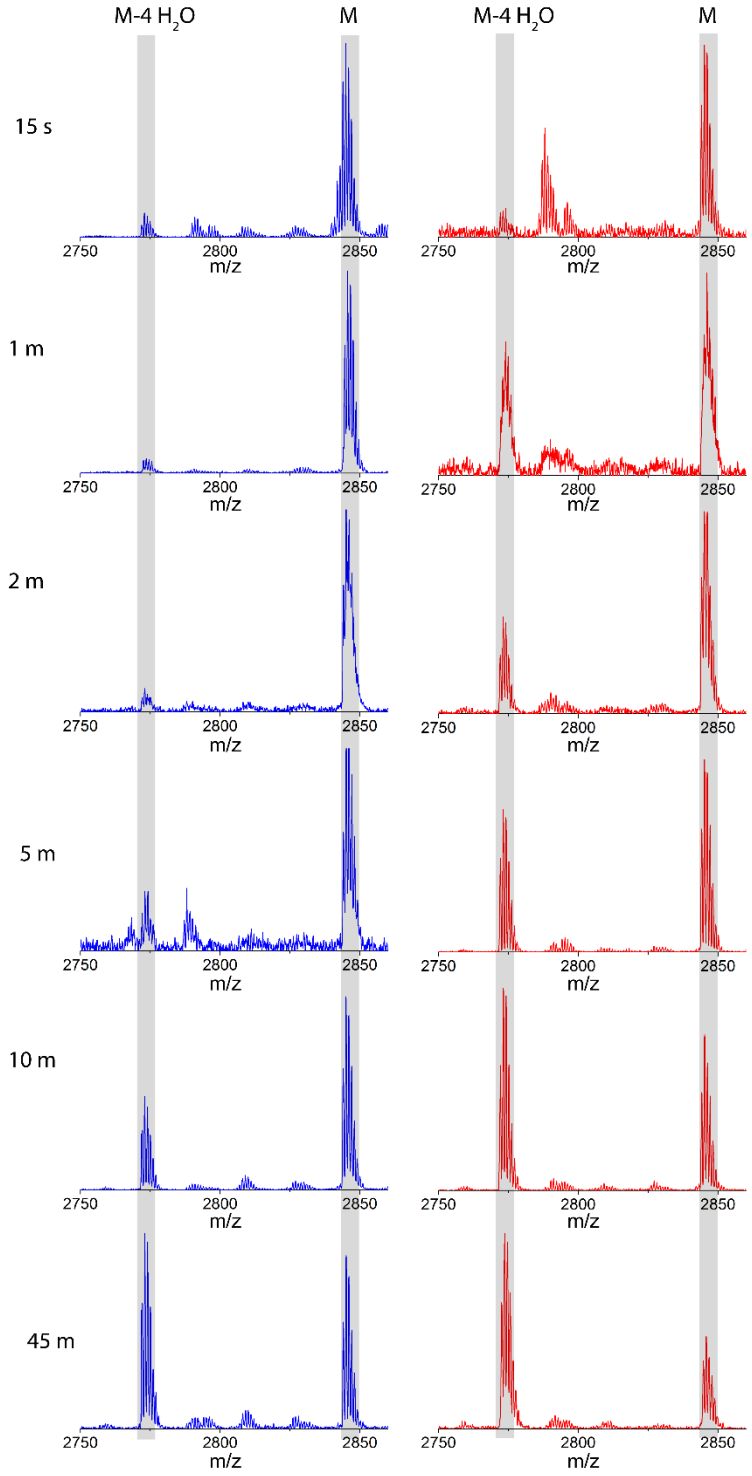

Figure 5.13. MALDI-TOF mass spectra of 25 μM WT LctA (blue) and LctA-34-28 (red) incubated with 1 μM LctM and treated with LysC to remove the leader peptide. Expected masses: M - 4 H_2O (2772.1), M (2844.2); calculated masses (WT at 45 min): M - 4 H_2O (2772.0), M (2844.1).

5.3 DISCUSSION AND OUTLOOK

Yeast surface display has several advantages over other display systems.[18] First, the relative large yeast cells can be sorted by fluorescence activated cell sorting (FACS), which is an extremely efficient high-throughput screening method compared to other traditional sorting methods such as 96-well plate sorting. Moreover, the application of FACS analysis also allows quantitative comparisons between mutants. In addition, ligand-binding affinity to target molecules may be directly measured on the yeast cell surface. Yeast surface display has been successfully applied to the engineering of many bioreceptor proteins including antibodies.[18, 19] In this chapter, yeast surface display was applied to engineer lanthipeptide leader peptides for tighter binding to their modification enzymes. The engineered LctA leader peptide (LctALea-34-28) showed 4-5 fold increase in the binding affinity compared to WT leader peptide in all FP and competitive FP experiments demonstrated in this chapter, which shows that this technique is suitable for the engineering of lanthipeptide leader peptides.

Previously, the engineering efforts of lanthipeptides mainly utilized site-directed mutagenesis and several studies generated lanthipeptide variants with enhanced activities.[20, 21] While site-directed mutagenesis allows the incorporation of specific mutations in the peptide, it suffers from several shortcomings: 1. mutagenesis of individual residues is time and effort-consuming, and in most cases, only part of the peptide sequence is covered; 2. the characterization of individual analogs is very low-throughput; 3. with limited knowledge about the peptide-enzyme binding in lanthipeptide biosynthesis, mutations are usually introduced randomly. In this chapter, a high-throughput engineering and screening strategy was applied to lanthipeptides, which could potentially solve these problems for the leader peptide. As

discussed in section 5.1, the generation of tighter-binding leader peptide could be applied in many aspects of lanthipeptide studies. Yeast surface display could potentially also be applied to engineer core peptide for desired properties, to tune the functions of lanthionine synthetases, or to generate a cyclic peptide library to screen for target functions such as disrupt specific protein-protein interactions.[22-24]

In the two rounds of sorting of the LctA yeast library, four mutations were found in all isolated yeast colonies with increased fluorescent signal. Three of these mutations were the substitution of a neutral amino acids with aspartic acid (N5D, N8D, and S16D), and one of these mutations (Q4L) was mutating a polar residue with the hydrophobic residue leucine. In addition, the only positively charged residue (Lys at the second position) was mutated to a neutral residue for all the mutants. The mutations were also inspected at nucleotide level, and biased mutations were observed. The leader peptides of lanthipeptides are naturally enriched with negatively charged residues,[3] and the accumulation of three conserved aspartic residues in the leader region is unlikely a coincidence. It is likely that negatively charged residues are beneficial to the leader-enzyme binding. Although from the NisB-NisA co-crystal structure,[18] only one aspartic acid in the leader peptide interacts with NisB, these negatively charged residues might contribute to the binding indirectly. More structural and biochemistry studies are needed to support this hypothesis.

The WT LctA peptide suffers from the problem of forming soluble aggregation and difficulty to ionize in the mass spectrometry studies, thus it is difficult to study the changes in the catalytic efficiency of LctM on the engineered LctA without the interference of these factors. However, the engineered LctA might have either increased product formation rate, or resulted

in less peptide aggregation, or very likely both. It can be concluded that by using yeast surface display I have generated a more efficient substrate for LctM, although the reasons for the improved product formation rate remain unclear. It is of interest to study the catalytic rates of LctM on substrates with different binding affinity, because this study could provide us with information about the rate limiting steps in lanthipeptide modifications.

There is room for improvements in this project in several aspects: First, in order to inspect the relationships between the catalytic rates of LctM and substrate binding affinity, it is required to eliminate the aggregation of the substrates. Methods described in this chapter including adding solubility tags were not successful in solving this problem. To solve this problem efficiently, the region of LctA peptides which causes aggregation needs to be investigated to make stable LctA peptide variants. Second, the application of engineered LctA leader peptides should be further explored. As mentioned in the introduction, a tighter binding leader peptide is useful in obtaining co-crystal structures with the synthetases. However, as no crystal structure was obtained for LctM, no crystallographic studies were carried out using the engineered peptide. A tighter binding leader peptide might lock the synthetase in the active state and facilitate the crystallization process, however, the 4-5 fold increase in binding affinity might not be enough for that purpose. The efforts to further engineer LctA leader peptide were unsuccessful due to the high frequency of ORF (open reading frame) shifts, but it is still possible that leader peptides with higher binding could be generated using LctA-34-28 as template. Third, the yeast surface display system described in this chapter could be applied to engineer other lanthipeptides for tighter binding affinity to the synthetases. With the peptide substrates which have no solubility and aggregation problems, the relationships between

catalytic rates and binding affinity could be inspected without complications. These experiments might also be informative in terms of enzymatic mechanisms of lanthionine syntheses and the evolution of lanthipeptide leader peptides. It is also possible to obtain a lanthionine synthetase which is locked in the active state by a tight binding leader. Such enzymes could be useful to provide a more efficient method for lanthipeptide production, which is especially important for the lanthipeptides of commercial interest.

5.4 EXPERIMENTAL PROCEDURES

5.4.1 Materials

Strains

Saccharomyces cerevisiae yeast display strain EBY100 (**a** *GAL1-AGA1*::*URA3 ura3-52 trp1 leu2D*1 *his3Δ200 pep4*::*HIS2 prb1*Δ*1.6R can1 GAL)* is commercially available from Invitrogen (cat. no. C839-00). In this study, the yeast strain was obtained from Professor David M. Kranz's lab.

E. coli strain used for cloning was DH10B™ (Invitrogen cat. no. 18290-015).

Yeast Display Plasmid

Yeast display plasmid pCT302 is commercially available from Invitrogen as pYD1 (cat. no. V835-01).

Restriction Enzymes

*Nhe*I (Invitrogen cat. no.15444-011), *Bgl*II (Invitrogen cat. no. 15213-010), *Xho*I (Invitrogen cat. no. 15231-012), and *Dpn*I (Invitrogen cat. no. 15242-019) were used as per the manufacturer's recommendations.

DNA Purification Kits

Plasmid rescue and purification kits used to purify plasmids from from *E. coli* were the QIAprep spin miniprep kit (QIAGEN cat. no. 27104). For plasmid purification from yeast the Zymoprep kit I or II (Zymoresearch cat. no. D2001 (I)) was used. QIAquick PCR purification kit (QIAGEN cat. no. 28104) was used to purify PCR products.

Yeast Media

The reecipe for YPD media for propagation of EBY100 was: 5 g yeast extract (BD Bacto cat. no. 212750), 10 g peptone (BD Bacto cat. no. 211677), and 10 g dextrose (EMD cat. no. 346351) dissolved in ddH_2O to give a final volume of 500 mL. YPD media was autoclaved and stored at room temperature. To make 1 sleeve of YPD plates, 7.5 g of agar was added to the above recipe.

The recipe for SD-citrate-CAA media for propagation of EBY100 transformed with pCT302 contained 7.4 g sodium citrate (Fisher cat. no. S279-500), 2.1 g citric acid monohydrate (Sigma cat. no. C1909), 2.5 g casamino acids (BD/Bacto cat. no. 223120), 10 g dextrose, and 3.35 g yeast nitrogen base without amino acids (Sigma cat. no. Y0626), dissolved in ddH_2O to make the final volume to be 500 mL. The dissolved media was filtered to sterilize and stored at 4 °C. For SD-citrate-CAA plates, 91 g sorbitol (Sigma cat. no. S1876), 7.5 g agar, 7.4 g sodium citrate, and 2.1 g citric acid monohydrate were dissolved in ddH_2O to make the final volume to be 400 mL, and autoclaved to sterilize. In a separate container, 2.5 g casamino acids, 10 g dextrose, and 3.35 g yeast nitrogen base without amino acids were dissolved in ddH_2O to make

the final volume to be 100 mL; then filtered to sterilize and added to the autoclaved solution when it was cool. The media was mixed thoroughly and poured into sterile petri dishes.

The SG-citrate-CAA media for induction of the yeast display construct was made as described above for 500 mL SD-citrate-CAA but substituting an equal mass of D-(+)-galactose (Sigma cat. no.G-0625) for dextrose.

Flow Cytometry/FACS Reagents

Phosphate-buffered saline (PBS): 1 X PBS solution was purchased from FISHER-Mediatech (cat. no. 21040CV). To make 100 mL PBS/0.5% BSA solution, 0.5 g bovine serum albumin (Sigma cat. no. A8531-1VL) was added to 1 X PBS solution and the solutions was filtered to sterilize.

Monoclonal antibodies were purchased: anti-c-myc chicken antibody (Molecular Probes cat. no. A-21281), Alexa488-labeled goat-anti-mouse secondary antibody (Invitrogen cat. no. A-11017), Alexa647-labeled goat-anti-chicken secondary antibody (Molecular Probes cat. no. A-21449), anti-flag M2 mouse antibody (Sigma cat. no. F3165).

5.4.2 Construction of yeast surface display plasmids for WT lanthipeptides

The lanthipeptide genes were PCR-amplified using Phusion polymerase by 30 cycles of denaturing (94 °C for 30 s), annealing (55 °C for 30 s), and extending (72 °C for 15 s) using primers and templates listed in Table 5.1. The PCR products were digested with NheI and XhoI restriction enzymes and ligated into the pCT302 vector to generate the construct pCT302-*LctA-myc/NisA-myc/ProcA3.2-myc* using a subcloning competent *E. coli* strain, such as DH10B. The sequence of the fused gene in the vector was confirmed using “Splice 4/L” and “T7 reverse” primers that flank the insert.

Table 5.1. Primers used in section 5.4.2.

Construct	Primer Name	Primer Sequences (5'-3')
pCT302-*LctA-myc* (Template: pET15b-*LctA*[15])	*LctA*_NheI_FP	GGTTCTGCTAGCATGAAAGAACAAAACTCT
	*LctA_myc*_XhoI_RP	TCAGATCTCGAGTTACAGATCTTCTTCAGAAATAAGTTTTTGTTCTTAAGAGCAGCAAGT
pCT302-*NisA-myc* (Template: pET15b-*NisA*[14])	*NisA*_NheI_FP	GCTAGCATGAGTACAAAAGATTTTAACTTGG
	*NisA_myc*_XhoI_RP	GCGTTCCTCGAGTTACAGATCTTCTTCAGAAATAAGTTTTTGTTCTTTGCTTACG
pCT302-*ProcAA3.2Lea-myc* (Template: pET15b-*procA3.2*[14])	*ProcA3.2*_NheI_FP	GGTTCTGCTAGCATGTCAGAAGAACAACTCAAGG
	*ProcA3.2Lea-myc*_ XhoI_RP	GGTTCTCTC GAG TTA CAG ATC TTC TTC AGA AAT AAG TTT TTG TTC TCC CCC AGC CAC ACC
Sequencing primers	Splice 4/L	GGCAGCCCCATAAACACACAGTAT
	T7 reverse	TAATACGACTCACTATAG

5.4.3 Error prone PCR

The PCR reaction mixture was prepared by adding 10 mM dNTPs, 0.3 ng/μL template DNA (gene of interest in pCT302 vector), primers listed in Table 5.2, 5 ng/μL Bovine Serum Albumin (BSA), 3.325 mM $MgCl_2$, 0.5 mM $MnCl_2$, 1:10 dilution of Taq Polymerase Reaction Buffer, 1 μL (5 unit) Taq Polymerase, and ddH_2O to take the final volume to be 100 μL. The target gene was PCR amplified by 30 cycles of denaturing (94 °C for 30s), annealing (55 °C for 30 s), and extending (72 °C for 30 s). The amplification was checked by running the PCR product on a 1% agarose gel. The PCR reaction was digested with 1 μL (20 unit) *Dpn*I for 1 h at 37°C to reduce insert background from remaining template. The reaction was then cleaned up by using a PCR Clean-up Kit from Qiagen, and eluted with 30 μL elution buffer.

Table 5.2. Primers used in section 5.4.3.

Primer Name	**Primer Sequences (5'-3')**
YSD_Error-prone_FP	GTTCTGGTGGTGGTGGTTCTGGTGGTGGTGG TTCTGCTAGC
YSD_Error-Prone_ LctA_RP	CAAAGTCGATTTTGTTACATCTACACTGTTGTTATCA GATCTCGAGTTACAGATCTTCTTCAGAAATAAGTTT TTGTTCAGAGCAGCAAGTAAATACAAATTGCCAGCT ATTCATATTACATTCATGAGAAATTGTATGAATAACT CCACTGCCGCCTTT
YSD_Error-Prone_ NisA_RP	CAAAGTCGATTTTGTTACATCTACACTGTTGTTATCA GATCTCGAGTTACAGATCTTCTTCAGAAATAAGTTT TTGTTCTTTGCTTACGTGAATACTACAATGACAAGT TGCTGTTTTCATGTTACAACCCATCAGAGCTCCTGT TTTACAACCGGGTGTACATAGCGAAATACTTGTAAT
YSD_Error-Prone_ ProcA3.2_RP	TACAGTGGGAACAAAGTCGATTTTGTTACATCTACA CTGTTGTTATCAGATCTCGAGTTACAGATCTTCTTCA GAAATAAGTTTTTGTTC

5.4.4 Preparation of pCT302 vector

Empty pCT302 vector was obtained from a Qiagen mini-prep and an aliquot of 50~80 μg vector was digested with NheI in appropriate buffer overnight at 37 °C. The digested reaction was cleaned up using the PCR Clean-up Kit from Qiagen. The NheI-digested vector was treated with XhoI in appropriate buffer for at least 1 h at 37 °C. The reaction was then cleaned up using the PCR Clean-up Kit. The NheI and XhoI digested vector was treated again with BglII in appropriate buffer for at least 1 h at 37 °C. Digesting with all three enzymes decreased the possibility of remaining full-length vector with the potential to re-circularize. The reaction was cleaned up using the PCR Clean-up Kit, and dephosphorylated using Calf Intestine Phosphatase (CIP) for 1 h at 37 °C in order to reduce vector background when transforming the yeast library. The final product was purified again using PCR Clean-up Kit.

5.4.5 Preparation of electrocompetent yeast for library construction

Two days before preparation of the library, 2 to 3 mL cultures of sterile YPD media were inoculated with a colony of EBY100 cells from a freshly-streaked YPD plate. The culture was grown to stationary phase (30 °C, 48 h, 220 rpm). The night before library transformation, a 2 L sterile flask containing 500 mL of YPD medium was inoculated with 5–9 mL of starter EBY100 culture. The culture was grown overnight at 30 °C with vigorous shaking until the cells reached an OD_{600} of 1.3–1.5. The culture was harvested in sterile centrifuge bottles that can hold 500 mL by spinning at 4,000× g at 4 °C. The pellet was resuspended vigorously in 80 mL of sterile ddH_2O. An aliquot of 10 mL of sterile 10× TE buffer (pH 7.5) was added to the yeast suspension and swirled to mix. An aliquot of 10 mL of 10× Lithium Acetate stock solution (1 M) was added to the yeast suspension and swirled to mix. The yeast suspension was then placed on an incubation shaker for 45 min at 30 °C with a speed of 100 rpm. After that, 2.5 mL of 1 M DTT was freshly prepared and added to the yeast suspension while swirling. The yeast suspension was then placed on an incubation shaker for 15 min at 30 °C with a speed of 100 rpm. The yeast suspension was diluted to 500 mL with sterile water. The cells were washed and concentrated three times by centrifuging at 4,000–6,000 × g, and pellets were resuspended successively with 250 mL of ice-cold water, 20–30 mL of ice-cold 1 M sorbitol, and 0.5 mL of ice-cold 1 M sorbitol. The final volume of resuspended, electrocompetent yeast cells was 1–1.5 mL.

5.4.6 Construction of EGFP-tagged enzymes

The *egfp* gene was amplified by PCR in 30 cycles of denaturing (98 °C for 30 s), annealing (55 °C for 30 s), and extending (1 min for each 1.5 kb) using the plasmid pCDNA3-EGFP

(obtained from Dr. Min Zeng) as templates and primers listed in Table 5.3. The PCR products were digested with NheI and BamHI restriction enzymes and ligated into pET28b vector treated with the same enzymes to generate the construct pET28b-N-*EGFP*. The sequences of the resulting products were confirmed by DNA sequencing.

Lanthionine synthetase genes were amplified by PCR using 30 cycles of denaturing (98 °C for 10 s), annealing (55 °C for 30 s), and extending (1 min for each 1.5 kb) using the primers listed in Table 5.3. The resulting DNA inserts and the pET28b-N-*EGFP* construct were digested with BamHI and XhoI and ligated. The sequences of the resulting products were confirmed by DNA sequencing.

Table 5.3. Primers used in section 5.4.6.

Construct	**Primer Name**	**Primer Sequences (5'-3')**
pET28b-N-*EGFP* (Template: pCDNA3-EGFP)	EGFP _NheI_FP	GGCTAGCATGGTGAGCAAGGGCGAGGA
	EGFP _BamI_RP	CGGATCCTCACTTGTACAGCTCGTCCAT
pET28b-N-*EGFP*-*LctM* (Template: pET28b-LctM[15], pET28b-N-*EGFP*)	LctM_BamI_FP	CATATGGGATCCATGAAAAAAAAGACTTAC
	LctM _XhoI_RP	GTGGTGCTCGAGTTAATCAACATATGG
pET28b-N-*EGFP*-*Nisc* (Template: pET15b-NisC[11], pET28b-N-*EGFP*)	NisC _BamI_FP	GGACGAGCTGTACAAGGGATCCATGAGGATAATGATGAA
	NisC _XhoI_RP	CGCTCGAGATCGCATTTCCTCTTCCCTCCTTT
pET28b-N-*EGFP*-*procM* (Template: pET28b-ProcM[14], pET28b-N-*EGFP*)	ProcM_EcoRI_FP	CAGGATCCGAATTCATGGAAAGTCCATCATCTTGG
	ProcM _NotI_RP	AAGGAAAAAAGCGGCCGCTTATTCAGTAGGCCAGAGA

5.4.7 Transformation and Characterization of Library

A mixture of error-prone PCR insert and linearized vector DNA was prepared at approximately a 6:1 molar ratio. In addition to the error-prone library mixture, additional controls were prepared to transform yeast cells: linearized vector alone (the same final concentration as in the library DNA mixture), error-prone insert alone (the same final concentration as in the library DNA mixture), unmutated gene in pCT302, and no DNA in the transformation. In a sterile, ice-cold Eppendorf tube, 40 μL of concentrated, electrocompetent yeast cells was mixed with 10–100 ng DNA (≤5 μL). The yeast cells were transferred to an ice-cold disposable electroporation cuvette (0.2 cm gap) and pulsed at 1.5 kV, 25 μF, 200 Ω. The time constant reported varied between 4.2 and 4.9 ms. An aliquot of 1 mL of ice-cold sorbitol was added to the cuvette to recover the yeast cells with gentle pipetting. The recovered yeast cell library from the experimental tube was collected in a single, sterile conical tube, while the controls were kept in separate tubes. An aliquot of 10 μL cells from each control transformation was spread onto a SD citrate-CAA plate. A 10 μL aliquot of undiluted, 100-fold diluted, or 500-fold diluted yeast cell library obtained above was plated onto SD-citrate-CAA plates to determine the size of the library. The plates were grown at 30°C for 48 to 72 h. The positive control transformation (pCT302 construct with WT lanthipeptide gene) was grown in several 3 mL culture tubes of SD-citrate-CAA media supplemented with 100 μg/mL ampicillin. The yeast cell library was allowed to recover by culturing in 500 mL SD-citrate-CAA media supplemented with 100 μg/mL ampicillin. The yeast cells were incubated at 30 °C with shaking at 220 rpm for 48 h. An aliquot of 10–50 mL of the expanded library was used to seed 1,000 mL SD-citrate-CAA media with 100 μg/mL ampicillin. The culture was incubated at 30 °C with shaking at 220 rpm

for another 24 h to reduce the number of untransformed or unhealthy cells. After the second passage of the library, 10–30 aliquots of the library containing 1 mL of dense culture (approximately 10^8 cells/mL) were mixed with 70 μL of DMSO and flash frozen. The remaining passaged library was stored in 3 mL aliquots in sterile culture tubes at 4 °C.

To analyze the library for diversity, plasmids were isolated from an aliquot of 0.5 mL of dense yeast cells from the library using the Zymoprep II kit. The isolated DNA was used transform *E. coli* DH10B cells. About 8–10 colonies were cultured. Plasmids were isolated from these cultures, and sent for sequencing using the "Splice 4/L" primer (Table 5.1).

5.4.8 Analyzing and Sorting by FACS

An aliquot of 1 mL of freshly passaged, dense library culture in SD-citrate-CAA media (about 10^8 yeast) was centrifuged in a sterile Eppendorf tube for 3 min at 3,000 rpm (800×g) in a tabletop microcentrifuge. The supernatant was removed by pipetting. The yeast cells were resuspended in SG-citrate-CAA media (galactose was substituted with glucose), and spun down again for 3 min at 3,000 rpm (800×g). The yeast cell pellet was resuspended in SG-citrate-CAA media containing 100 μg/mL ampicillin. This culture was incubated at 20 °C with shaking at 220 rpm for 24–48 h to induce the yeast cells to express AGA2- lanthipeptides. After the induction, the yeast was stored at 4 °C in SG-citrate-CAA media for up to 2 weeks.

The cells were stained for flow cytometry or sorting. First, the appropriate number of cells in SG-citrate-CAA media was centrifuged at 3,000 rpm (800×g) for 3 min. The supernatant was discarded, and pellet was rinsed one time with PBS/0.5% BSA. The cells were resuspended in an appropriate volume of PBS/0.5% BSA containing the desired primary staining reagent. Large samples for sorting by FACS were stained in a large enough volume to stay in suspension

(often 0.5–1.0 mL for 10^8 cells). Smaller, test samples (0.5–1.0 × 10^6 cells) were stained in 20–50 μL of PBS/0.5% BSA. The yeast cells were incubated in the staining solution at room temperature for 30 min or on ice for 45 min. The yeast cells were gently agitated several times to make sure that they remain in suspension. The cells were centrifuged at 3,000 rpm (800×g) for 3 min, and washed with a large excess (20 volumes) ice-cold of PBS/0.5% BSA. All the steps from this point forward were performed on ice. The yeast cells were resuspended in the appropriate volume of ice-cold secondary staining solution (fluorescently labeled reagents in PBS/0.5% BSA). Samples were incubated on ice in the dark for 30–45 min. Control samples were incubated with secondary, but not primary stain to determine the level of nonspecific binding of the secondary reagent. To stain yeast cells with EGFP-tagged enzymes, only one round of staining and washing was performed at room temperature.

The cells were spun down at 3,000 rpm (800×g) for 3 min, washed with 1 mL of ice-cold PBS/0.5% BSA, and resuspended to an appropriate volume of ice-cold PBS/0.5% BSA for flow cytometry analysis or sorting. The samples were run on the flow cytometer or FACS instrument, using a gate on the forward scatter vs. side scatter plot to focus only on the live, healthy cells. Within the live population, a second gate was set up based on fluorescence that collects the yeast cells with the highest expression of the fluorescent probe of interest. In the first round of sorting, about 5% of the most positive yeast cells with enhanced fluorescence signal were collected. Smaller fractions of the yeast cells were collected on subsequent sorts, down to 0.1–1%. The yeast cells inside of both gates (scatter and fluorescence) were sorted into a sterile cell culture tube on ice. Usually at least 50,000 cells were collected. The collected yeast cells were diluted into SD-citrate-CAA media with ampicillin and transferred to a culture

tube and incubated at 30 °C with shaking at 220 rpm. The resulting cultures were used for the next round of sorting process.

The library sorting and expansion procedures were repeated until a distinctly positive population in the expanded library appeared. The yeast cells from that population were isolated and used for plasmid extraction. The obtained plasmids were used to transform fresh yeast cells. The yeast colonies with individual plasmids were cultured and analyzed by FACS to measure the mean fluorescence intensity after staining the EGFP-tagged enzymes. The best hit was used as template for the next round of library construction and sorting. The overall process is shown in Figure 5.15.

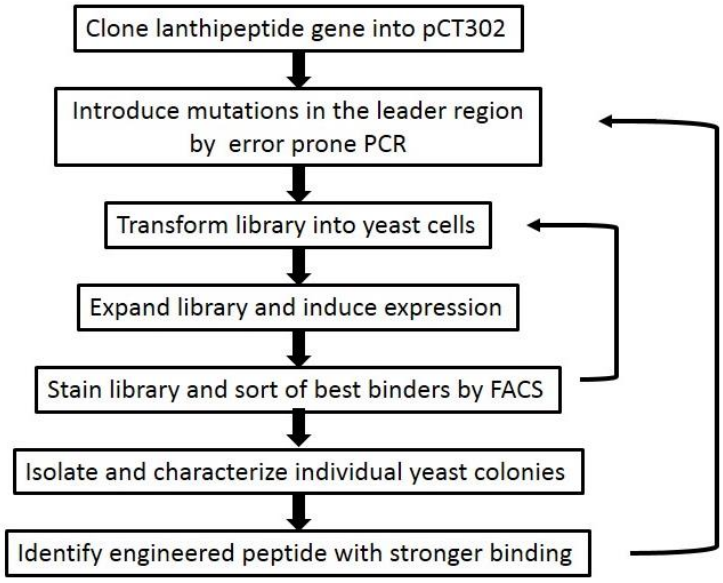

Figure 5.15. Flow chart showing the sorting process of yeast cell library.

5.4.9 Construction of ConFusion LctM enzyme

The pET28b LctCE-$(GS)_{15}$ plasmid was used as template.[25] Subcloning was carried out to displace the WT LctA leader peptide with an engineered leader peptide using primers listed in Table 5.4.

Table 5.4. Primers used in section 5.4.9.

Construct	Primer Name	Primer Sequences (5'-3')
LctCE-$(GS)_{15}$-34 (Template: pCT302-LctA-34)	LctA_NheI_FP	CGCGGCACGGCTAGCATGAAAG AACAAAAC
	LctALea_SpeI_RP	CCAGAACCACTAGTTGCACCTAA AATAAGTCCAA
LctCE-$(GS)_{15}$-34-28 (Template: pCT302-LctA-34-28)	LctA-34-28_NheI_FP	CGCGGCACGGCTAGCATGAATG AACTAGAC
	LctALea_SpeI_RP	CCAGAACCACTAGTTGCACCTAA AATAAGTCCAA

5.4.10 Construction of LctA/LctALea expression plasmids

The peptide genes were amplified by PCR in 30 cycles of denaturing (98 °C for 30 s), annealing (55 °C for 30 s), and extending (72 °C for 15 s) using the plasmids isolated from yeast surface display as templates and primers listed in Table 5.5. The PCR products were digested with the indicated restriction enzymes and ligated into vectors listed in Table 5.5. The sequences of the resulting products were confirmed by DNA sequencing.

Table 5.5. Primers used in section 5.4.10.

Construct	Primer Name	Primer Sequences (5'-3')
pET15b-LctA-34 (template: pCT302-LctA-34, vector: pET15b)	LctA-34_NdeI_FP	GGTTCTCATATGGGAAAAGAAC AAAACTCATT TG
	LctA_XhoI_RP	GCCCGCGGATCCTTAAGAGCA GCAAGTA
pET15b-LctA-34-28 (template: pCT302-LctA-34-28 vector: pET15b)	LctA-34-28_NdeI_FP	GGTTCTCATATGGGAAATGAAC TAGACTCTTTTG
	LctA_XhoI_RP	GCCCGCGGATCCTTAAGAGCA GCAAGTA
pET15b-MBP-LctA (template: pET15b-LctA vector: pET28b-MBP*)	LctA_BamHI_FP	AAAAAAGGATCCATGAAAGAA CAAAACTC
	LctA_XhoI_RP	GCCCGCGGATCCTTAAGAGCA GCAAGTA

(Table 5.5 continued)

pET15b-MBP-LctA-34-28 (template: pCT302-LctA-34-28 vector: pET28b-MBP*)	LctA-34-28_ BamHI_FP	AAAAAAGGATCCATGAATGAACTAG ACTC
	LctA_XhoI_ RP	GCCCGCGGATCCTTAAGAGCAGCAA GTA
pET15b-MBP-LctALea (template: pET15b-LctA vector: pET28b-MBP*)	LctA_BamHI_ FP	AAAAAAGGATCCATGAAAGAACAAA ACTC
	LctALea_ XhoI_RP	CGGATCCTCGAGTTATGCACCTAAAA TAAGGTC
pET15b-MBP-LctALea-34-28 (template: pCT302-LctA-34-28 vector:pET28b-MBP*)	LctA-34-28_ BamHI_FP	AAAAAAGGATCCATGAATGAACTAG ACTC
	LctALea_ XhoI_RP	CGGATCCTCGAGTTATGCACCTAAAA TAAGGTC
pET-SUMO-LctA (template: pET15b-LctA vector: pET-SUMO*)	LctA_BamHI_ FP	AAAAAAGGATCCATGAAAGAACAAA ACTC
	LctA_NotI_RP	GGAAAAAAGCGGCCGCTTAAGAGC AGCAAGTA
pET-SUMO-LctA-34-28 (template: pCT302-LctA-34-28 vector: pET-SUMO*)	LctA-34-28_ BamHI_FP	AAAAAAGGATCCATGAATGAACTAG ACTC
	LctA_NotI_RP	GGAAAAAAGCGGCCGCTTAAGAGC AGCAAGTA
pET15b-MBP-LctA-1-38 (template: pET15b-LctA vector: pET28b-MBP*)	LctA_BamHI_ FP	AAAAAAGGATCCATGAAAGAACAAA ACTC
	LctA1-38_ XhoI_ RP	AAAAAACTCGAGCTAACATTCATGA GAAATTG
pET15b-MBP-LctA-34-28-1-38 (template: pCT302-LctA-34-28 vector:pET28b-MBP*)	LctA-34-28_ BamHI_FP	AAAAAAGGATCCATGAATGAACTAG ACTC
	LctA1-38_ XhoI_ RP	AAAAAACTCGAGCTAACATTCATGA GAAATTG

* The plasmid pET28b-MBP was obtained from Manny Ortega, pET-SUMO was obtained from Yue Hao.

5.4.11 Labelling LctA leader peptides with fluorescein isothiocyanate (FITC)

FITC (purchased from Sigma) was dissolved in DMF at 10 mg/mL. An aliquot of 1 mg peptide was dissolved in 0.5 mL of conjugation buffer (50 mM borate buffer, pH 8.5). An excess amount of FITC was added (FITC concentration was 20-fold that of the peptide concentration) to the peptide solution and the mixture was incubated at room temperature in the dark for 1~3 h. The resulting peptide mixture was subjected to analytical HPLC analysis and the labelled peptide was purified.

5.4.12 Analytical HPLC analysis

Samples were run on an Agilent analytical HPLC system equipped with a Thermo Scientific Gold C18 column. The solvent used was 98% solvent A and 2% solvent B for 1 min, followed by an increase to 100% solvent B over 45 min with a flow rate of 1.0 mL/min (solvent A: 80% acetonitrile with 0.086% TFA, solvent B: ddH_2O with 0.1% TFA).

5.4.13 Fluorescence polarization (FP) assay

All FP experiments were performed on a Synergy H4 Hybrid plate reader (BioTek) using a 96-well black opaque half area plate (Costar 3694). Increasing amounts of LctM enzyme were added to fluorescein-labeled peptides (5 or 20 nM) in LanM Start buffer (20 mM HEPES, 1 M NaCl, pH 7.5). Data analysis was performed using Origin 9.0, and curves were fitted to a hyperbolic equation to calculate K_D ($y = (a*x)/(K_D + x) + c$).Competitive FP assays were completed with LctM (1.7 μM) and a fluorescein-labeled peptide (5 or 20 nM) in LanM Start buffer. Increasing amounts of unlabeled peptide were added to displace the labeled peptide.

5.4.14 *In vitro* reconstitution of the activity of LctM and ConFusion enzymes

The *in vitro* enzymatic assay included: 1 mM ATP, 10 mM DTT, 50 mM HEPES (pH 7.5), 25 μg/mL bovine serum albumin, 10 mM $MgCl_2$ and various amounts of peptide and enzyme indicated in the figure legends. The *in trans* assay included: LctA core peptide (final concentration 20 μM), purified His_6-LctCE (final concentration 2 μM), 10 mM $MgCl_2$, 2 mM ATP, 10 mM DTT, 25 μg/mL bovine serum albumin and 50 mM HEPES buffer (pH 7.5). The extent of reaction was monitored by periodically removing aliquots of the reaction, which were quenched by the addition of formic acid to 0.5% final concentration (pH 1-2), desalted by ZipTipC18 and analyzed by downstream experiments.

5.4.15 Experimental procedures listed in other chapters

Expression and purification of unmodified and modified peptides (section 2.4.5);

Expression and purification of LanM enzymes (section 2.4.6);

Bioactivity assay (section 2.4.8);

Iodoacetamide assay (section 2.4.9);

ESI-LC-MS experiments (section 2.4.10).

5.5 REFERENCES

1. Knerr, P. J., and van der Donk, W. A. (2012) Discovery, biosynthesis, and engineering of lantipeptides, *Annu. Rev. Biochem. 81*, 479-505.

2. Chatterjee, C., Paul, M., Xie, L., and van der Donk, W. A. (2005) Biosynthesis and mode of action of lantibiotics, *Chem. Rev. 105*, 633-684.

3. Oman, T. J., and van der Donk, W. A. (2010) Follow the leader: the use of leader peptides to guide natural product biosynthesis, *Nat. Chem. Biol. 6*, 9-18.

4. Kluskens, L. D., Kuipers, A., Rink, R., de Boef, E., Fekken, S., Driessen, A. J., Kuipers, O. P., and Moll, G. N. (2005) Post-translational Modification of Therapeutic Peptides By NisB, the Dehydratase of the Lantibiotic Nisin, *Biochemistry 44*, 12827-12834.

5. Chatterjee, C., Patton, G. C., Cooper, L., Paul, M., and van der Donk, W. A. (2006) Engineering dehydro amino acids and thioethers into peptides using lacticin 481 synthetase, *Chem. Biol. 13*, 1109-1117.

6. Levengood, M. R., and van der Donk, W. A. (2008) Use of Lantibiotic Synthetases for the Preparation of Bioactive Constrained Peptides, *Bioorg. Med. Chem. Lett. 18*, 3025-3028.

7. Oman, T. J., Knerr, P. J., Bindman, N. A., Velasquez, J. E., and van der Donk, W. A. (2012) An engineered lantibiotic synthetase that does not require a leader peptide on its substrate, *J. Am. Chem. Soc. 134*, 6952–6955.

8. Kuipers, O. P., Bierbaum, G., Ottenwalder, B., Dodd, H. M., Horn, N., Metzger, J., Kupke, T., Gnau, V., Bongers, R., van den Bogaard, P., Kosters, H., Rollema, H. S., de Vos, W. M., Siezen, R. J., Jung, G., Gotz, F., Sahl, H. G., and Gasson, M. J. (1996) Protein engineering of lantibiotics, *Antonie van Leeuwenhoek 69*, 161-169.

9. Ortega, M. A., Hao, Y., Zhang, Q., Walker, M. C., van der Donk, W. A., and Nair, S. K. (2014) Structure and mechanism of the tRNA-dependent lantibiotic dehydratase NisB *Nature 517*, 509-512.

10. Li, B., Yu, J. P., Brunzelle, J. S., Moll, G. N., van der Donk, W. A., and Nair, S. K. (2006) Structure and mechanism of the lantibiotic cyclase involved in nisin biosynthesis, *Science 311*, 1464-1467.

11. Jungmann, N. A., Krawczyk, B., Tietzmann, M., Ensle, P., and Süssmuth, R. D. (2014) Dissecting Reactions of Nonlinear Precursor Peptide Processing of the Class III Lanthipeptide Curvopeptin, *J. Am. Chem. Soc. 136*, 15222-15228.

12. Zhang, Q., Yu, Y., Velasquez, J. E., and van der Donk, W. A. (2012) Evolution of lanthipeptide synthetases, *Proc. Natl. Acad. Sci. U.S.A. 109*, 18361-18366.

13. Richman, S. A., Kranz, D. M., and Stone, J. D. (2009) Biosensor detection systems: engineering stable, high-affinity bioreceptors by yeast surface display, *Methods Mol. Biol. 504*, 323-350.

14. Li, B., Sher, D., Kelly, L., Shi, Y., Huang, K., Knerr, P. J., Joewono, I., Rusch, D., Chisholm, S. W., and van der Donk, W. A. (2010) Catalytic promiscuity in the biosynthesis of cyclic peptide secondary metabolites in planktonic marine cyanobacteria, *Proc. Natl. Acad. Sci. U.S.A. 107*, 10430-10435.

15. Xie, L., Miller, L. M., Chatterjee, C., Averin, O., Kelleher, N. L., and van der Donk, W. A. (2004) Lacticin 481: in vitro reconstitution of lantibiotic synthetase activity, *Science 303*, 679-681.

16. Lee, M. V., Ihnken, L. A., You, Y. O., McClerren, A. L., van der Donk, W. A., and Kelleher, N. L. (2009) Distributive and directional behavior of lantibiotic synthetases revealed by high-resolution tandem mass spectrometry, *J. Am. Chem. Soc. 131*, 12258-12264.

17. Yu, Y., Mukherjee, S., and van der Donk, W. A. (2015) Product Formation by the Promiscuous Lanthipeptide Synthetase ProcM is under Kinetic Control, *J. Am. Chem. Soc. 137*, 5140-5148.

18. Kondo, A., and Ueda, M. (2004) Yeast cell-surface display--applications of molecular display, *Appl. Microbiol. Biotechnol. 64*, 28-40.

19. Richman, S. A., and Kranz, D. M. (2007) Display, engineering, and applications of antigen-specific T cell receptors, *Biomol. Eng. 24*, 361-373.

20. Liu, W., and Hansen, J. N. (1992) Enhancement of the chemical and antimicrobial properties of subtilin by site-directed mutagenesis, *J. Biol. Chem. 267*, 25078-25085.

21. Field, D., Connor, P. M., Cotter, P. D., Hill, C., and Ross, R. P. (2008) The generation of nisin variants with enhanced activity against specific gram-positive pathogens, *Mol. Microbiol. 69*, 218-230.

22. Shivange, A. V., and Daugherty, P. S. (2015) De novo discovery of bioactive cyclic peptides using bacterial display and flow cytometry, *Methods Mol. Biol. 1248*, 139-153.

23. Zhang, Z., Lin, Z., Zhou, Z., Shen, H. C., Yan, S. F., Mayweg, A. V., Xu, Z., Qin, N., Wong, J. C., Zhang, Z., Rong, Y., Fry, D. C., and Hu, T. (2014) Structure-Based Design and Synthesis of Potent Cyclic Peptides Inhibiting the YAP-TEAD Protein-Protein Interaction, *ACS Med. Chem. Lett. 5*, 993-998.

24. Glas, A., Bier, D., Hahne, G., Rademacher, C., Ottmann, C., and Grossmann, T. N. (2014) Constrained peptides with target-adapted cross-links as inhibitors of a pathogenic protein-protein interaction, *Angew. Chem., Int. Ed. Engl. 53*, 2489-2493.

25. Goto, Y., Li, B., Claesen, J., Shi, Y., Bibb, M. J., and van der Donk, W. A. (2010) Discovery of unique lanthionine synthetases reveals new mechanistic and evolutionary insights, *PLoS Biol. 8*, e1000339.